



(11) EP 1 790 660 A2

## (12) EUROPEAN PATENT APPLICATION

(43) Date of publication:  
30.05.2007 Bulletin 2007/22(51) Int Cl.:  
*C07K 14/22 (2006.01)*

(21) Application number: 06076718.3

(22) Date of filing: 28.02.2001

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE TR

- Massignani, Vega  
53100 Siena (IT)
- Giuliani, Marzia Monica  
53100 Siena (IT)
- Pizza, Mariagrazia  
53100 Siena (IT)

(30) Priority: 28.02.2000 GB 0004695  
13.11.2000 GB 0027675(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:  
01914109.2 / 1 259 627(74) Representative: Marshall, Cameron John et al  
Carpmaels & Ransford  
43-45 Bloomsbury Square  
London WC1A 2RA (GB)(71) Applicant: Novartis Vaccines and Diagnostics  
S.r.l.  
53100 Siena SI (IT)Remarks:

- This application was filed on 13 - 09 - 2006 as a divisional application to the application mentioned under INID code 62.
- The sequence listing, which is published as annex to the application documents, was filed after the date of filing. The applicant has declared that it does not include matter which goes beyond the content of the application as filed.

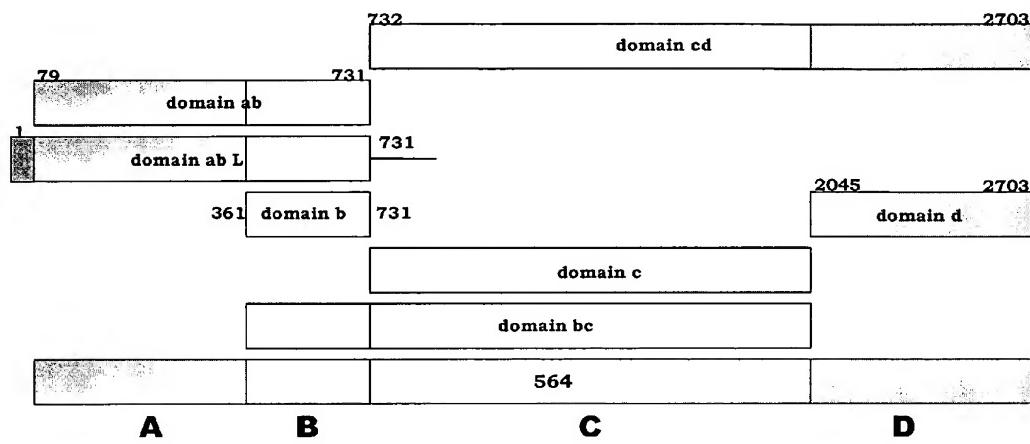
(72) Inventors:

- Arico, Maria B.  
53100 Siena (IT)
- Comanducci, Maurizio  
53100 Siena (IT)
- Galeotti, Cesira  
53100 Siena (IT)

## (54) Heterologous expression of neisserial proteins

(57) Alternative approaches to the heterologous expression of the proteins of *Neisseria meningitidis* and *Neisseria gonorrhoeae*. These approaches typically af-

fect the level of expression, the ease of purification, the cellular localisation, and/or the immunological properties of the expressed protein.

***FIGURE 8***

**Description****TECHNICAL FIELD**

5 [0001] This invention is in the field of protein expression. In particular, it relates to the heterologous expression of proteins from *Neisseria* (e.g. *N.gonorrhoeae* or, preferably, *N.meningitidis*).

**BACKGROUND ART**

10 [0002] International patent applications WO99/24578, WO99/36544, WO99/57280 and WO00/22430 disclose proteins from *Neisseria meningitidis* and *Neisseria gonorrhoeae*. These proteins are typically described as being expressed in *E.coli* (i. e. heterologous expression) as either N-terminal GST-fusions or C-terminal His-tag fusions, although other expression systems, including expression in native *Neisseria*, are also disclosed.

15 [0003] It is an object of the present invention to provide alternative and improved approaches for the heterologous expression of these proteins. These approaches will typically affect the level of expression, the ease of purification, the cellular localisation of expression, and/or the immunological properties of the expressed protein.

**DISCLOSURE OF THE INVENTION**20 ***Nomenclature herein***

[0004] The 2166 protein sequences disclosed in WO99/24578, WO99/36544 and WO99/57280 are referred to herein by the following SEQ# numbers:

25	Application	Protein sequences	SEQ# herein
	WO99/24578	Even SEQ IDs 2-892	SEQ#s 1-446
	WO99/36544	Even SEQ IDs 2-90	SEQ#s 447-491
30		Even SEQ IDs 2-3020	SEQ#s 492-2001
	WO99/57280	Even SEQ IDs 3040-3114 SEQ IDs 3115-3241	SEQ#s 2002-2039 SEQ#s 2040-2166

35 [0005] In addition to this SEQ# numbering, the naming conventions used in WO99/24578, WO99/36544 and WO99/57280 are also used (e.g. 'ORF4', 'ORF40', 'ORF40-1' etc. as used in WO99/24578 and WO99/36544; 'm919', 'g919' and 'a919' etc. as used in WO99/57280).

[0006] The 2160 proteins NMB0001 to NMB2160 from Tettelin et al. [Science (2000) 287:1809-1815] are referred to herein as SEQ#s 2167-4326 [see also WO00/66791].

40 [0007] The term 'protein of the invention' as used herein refers to a protein comprising:

- (a) one of sequences SEQ#s 1-4326; or
- (b) a sequence having sequence identity to one of SEQ#s 1-4326; or
- (c) a fragment of one of SEQ#s 1-4326.

45 [0008] The degree of 'sequence identity' referred to in (b) is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more). This includes mutants and allelic variants [e.g. see WO00/66741]. Identity is preferably determined by the Smith-Waterman homology search algorithm as implemented in the MPSRCH program (Oxford Molecular), using an affine gap search with parameters *gap open penalty*=12 and *gap extension penalty*=1. Typically, 50% identity or more between two proteins is considered to be an indication of functional equivalence.

50 [0009] The 'fragment' referred to in (c) should comprise at least n consecutive amino acids from one of SEQ#s 1-4326 and, depending on the particular sequence, n is 7 or more (eg. 8, 10, 12, 14, 16, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100 or more). Preferably the fragment comprises an epitope from one of SEQ#s 1-4326. Preferred fragments are those disclosed in WO00/71574 and WO01/04316.

[0010] Preferred proteins of the invention are found in *N.meningitidis* serogroup B.

55 [0011] Preferred proteins for use according to the invention are those of serogroup B *N.meningitidis* strain 2996 or strain 394/98 (a New Zealand strain). Unless otherwise stated, proteins mentioned herein are from *N.meningitidis* strain 2996. It will be appreciated, however, that the invention is not in general limited by strain. References to a particular

protein (e.g. '287', '919' etc.) may be taken to include that protein from any strain.

#### ***Non-fusion expression***

- 5 [0012] In a first approach to heterologous expression, no fusion partner is used, and the native leader peptide (if present) is used. This will typically prevent any 'interference' from fusion partners and may alter cellular localisation and/or post-translational modification and/or folding in the heterologous host.
- [0013] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) no fusion partner is used, and (b) the protein's native leader peptide (if present) is used.
- 10 [0014] The method will typically involve the step of preparing an vector for expressing a protein of the invention, such that the first expressed amino acid is the first amino acid (methionine) of said protein, and last expressed amino acid is the last amino acid of said protein (*i.e.* the codon preceding the native STOP codon).
- [0015] This approach is preferably used for the expression of the following proteins using the native leader peptide: 111, 149, 206, 225-1, 235, 247-1, 274, 283, 286, 292, 401, 406, 502-1, 503, 519-1, 525-1, 552, 556, 557, 570, 576-1, 15 580, 583, 664, 759, 907, 913, 920-1, 936-1, 953, 961, 983, 989, Orf4, Orf7-1, Orf9-1, Orf23, Orf25, Orf37, Orf38, Orf40, Orf40.1, Orf40.2, Orf72-1, Orf76-1, Orf85-2, Orf91, Orf97-1, Orf119, Orf143.1, NMB0109 and NMB2050. The suffix 'L' used herein in the name of a protein indicates expression in this manner using the native leader peptide.
- [0016] Proteins which are preferably expressed using this approach using no fusion partner and which have no native leader peptide include: 008, 105, 117-1, 121-1, 122-1, 128-1, 148, 216, 243, 308, 593, 652, 726, 926, 982, Orf83-1 and 20 Orf143-1.
- [0017] Advantageously, it is used for the expression of ORF25 or ORF40, resulting in a protein which induces better anti-bactericidal antibodies than GST- or His-fusions.
- [0018] This approach is particularly suited for expressing lipoproteins.

25 ***Leader-peptide substitution***

- [0019] In a second approach to heterologous expression, the native leader peptide of a protein of the invention is replaced by that of a different protein. In addition, it is preferred that no fusion partner is used. Whilst using a protein's own leader peptide in heterologous hosts can often localise the protein to its 'natural' cellular location, in some cases the leader sequence is not efficiently recognised by the heterologous host. In such cases, a leader peptide known to drive protein targeting efficiently can be used instead.
- [0020] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) the protein's leader peptide is replaced by the leader peptide from a different protein and, optionally, (b) no fusion partner is used.
- 35 [0021] The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; manipulating said nucleic acid to remove nucleotides that encode the protein's leader peptide and to introduce nucleotides that encode a different protein's leader peptide. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector. The expressed protein will consist of the replacement leader peptide at the N-terminus, followed by the protein of the invention minus its leader peptide.
- 40 [0022] The leader peptide is preferably from another protein of the invention (*e.g.* one of SEQ#s 1-4326), but may also be from an *E.coli* protein (*e.g.* the OmpA leader peptide) or an *Erwinia carotovora* protein (*e.g.* the PelB leader peptide), for instance.
- [0023] A particularly useful replacement leader peptide is that of ORF4. This leader is able to direct lipidation in *E.coli*, improving cellular localisation, and is particularly useful for the expression of proteins 287, 919 and ΔG287. The leader peptide and N-terminal domains of 961 are also particularly useful.
- 45 [0024] Another useful replacement leader peptide is that of *E.coli* OmpA. This leader is able to direct membrane localisation of *E.coli*. It is particularly advantageous for the expression of ORF1, resulting in a protein which induces better anti-bactericidal antibodies than both fusions and protein expressed from its own leader peptide.
- 50 [0025] Another useful replacement leader peptide is MKKYLFSAA. This can direct secretion into culture medium, and is extremely short and active. The use of this leader peptide is not restricted to the expression of Neisserial proteins - it may be used to direct the expression of any protein (particularly bacterial proteins).

#### ***Leader-peptide deletion***

- 55 [0026] In a third approach to heterologous expression, the native leader peptide of a protein of the invention is deleted. In addition, it is preferred that no fusion partner is used.
- [0027] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) the protein's leader peptide is deleted and, optionally, (b) no fusion partner is used.

[0028] The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; manipulating said nucleic acid to remove nucleotides that encode the protein's leader peptide. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector. The first amino acid of the expressed protein will be that of the mature native protein.

5 [0029] This method can increase the levels of expression. For protein 919, for example, expression levels in *E.coli* are much higher when the leader peptide is deleted. Increased expression may be due to altered localisation in the absence of the leader peptide.

[0030] The method is preferably used for the expression of 919, ORF46, 961, 050-1, 760 and 287.

#### 10 ***Domain-based expression***

[0031] In a fourth approach to heterologous expression, the protein is expressed as domains. This may be used in association with fusion systems (e.g. GST or His-tag fusions).

15 [0032] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) at least one domain in the protein is deleted and, optionally, (b) no fusion partner is used.

[0033] The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; manipulating said nucleic acid to remove at least one domain from within the protein. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector. Where no fusion partners are used, the first amino acid of the expressed protein will be that of a domain of the protein.

20 [0034] A protein is typically divided into notional domains by aligning it with known sequences in databases and then determining regions of the protein which show different alignment patterns from each other.

[0035] The method is preferably used for the expression of protein 287. This protein can be notionally split into three domains, referred to as A B & C (see Figure 5). Domain B aligns strongly with IgA proteases, domain C aligns strongly with transferrin-binding proteins, and domain A shows no strong alignment with database sequences. An alignment of polymorphic forms of 287 is disclosed in WO00/66741.

25 [0036] Once a protein has been divided into domains, these can be (a) expressed singly (b) deleted from with the protein e.g. protein ABCD → ABD, ACD, BCD etc. or (c) rearranged e.g. protein ABC → ACB, CAB etc. These three strategies can be combined with fusion partners is desired.

30 [0037] ORF46 has also been notionally split into two domains - a first domain (amino acids 1-433) which is well-conserved between species and serogroups, and a second domain (amino acids 433-608) which is not well-conserved. The second domain is preferably deleted. An alignment of polymorphic forms of ORF46 is disclosed in WO00/66741.

[0038] Protein 564 has also been split into domains (Figure 8), as have protein 961 (Figure 12) and protein 502 (amino acids 28-167 of the MC58 protein).

#### 35 ***Hybrid proteins***

[0039] In a fifth approach to heterologous expression, two or more (e.g. 3, 4, 5, 6 or more) proteins of the invention are expressed as a single hybrid protein. It is preferred that no non-Neisserial fusion partner (e.g. GST or poly-His) is used.

40 [0040] This offers two advantages. Firstly, a protein that may be unstable or poorly expressed on its own can be assisted by adding a suitable hybrid partner that overcomes the problem. Secondly, commercial manufacture is simplified - only one expression and purification need be employed in order to produce two separately-useful proteins.

[0041] Thus the invention provides a method for the simultaneous heterologous expression of two or more proteins of the invention, in which said two or more proteins of the invention are fused (i.e. they are translated as a single polypeptide chain).

45 [0042] The method will typically involve the steps of: obtaining a first nucleic acid encoding a first protein of the invention; obtaining a second nucleic acid encoding a second protein of the invention; ligating the first and second nucleic acids. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector.

[0043] Preferably, the constituent proteins in a hybrid protein according to the invention will be from the same strain.

50 [0044] The fused proteins in the hybrid may be joined directly, or may be joined via a linker peptide e.g. via a poly-glycine linker (*i.e.* G<sub>n</sub> where n = 3, 4, 5, 6, 7, 8, 9, 10 or more) or via a short peptide sequence which facilitates cloning. It is evidently preferred not to join a ΔG protein to the C-terminus of a poly-glycine linker.

[0045] The fused proteins may lack native leader peptides or may include the leader peptide sequence of the N-terminal fusion partner.

[0046] The method is well suited to the expression of proteins orf1, orf4, orf25, orf40, Orf46/46.1, orf83, 233, 287, 292L, 564, 687, 741, 907, 919, 953, 961 and 983.

[0047] The 42 hybrids indicated by 'X' in the following table of form NH<sub>2</sub>-A-B-COOH are preferred:

<b>↓A</b>	<b>B→</b>	<b>ORF46.1</b>	<b>287</b>	<b>741</b>	<b>919</b>	<b>953</b>	<b>961</b>	<b>983</b>
	<b>ORF46.1</b>		<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
5	<b>287</b>	<b>X</b>		<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
	<b>741</b>	<b>X</b>	<b>X</b>		<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>
10	<b>919</b>	<b>X</b>	<b>X</b>	<b>X</b>		<b>X</b>	<b>X</b>	<b>X</b>
	<b>953</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>		<b>X</b>	<b>X</b>
	<b>961</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>		<b>X</b>
	<b>983</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	

15 [0048] Preferred proteins to be expressed as hybrids are thus ORF46.1, 287, 741, 919, 953, 961 and 983. These may be used in their essentially full-length form, or poly-glycine deletions ( $\Delta G$ ) forms may be used (e.g.  $\Delta G$ -287,  $\Delta GTbp2$ ,  $\Delta G741$ ,  $\Delta G983$  etc.), or truncated forms may be used (e.g.  $\Delta 1$ -287,  $\Delta 2$ -287 etc.), or domain-deleted versions may be used (e.g. 287B, 287C, 287BC, ORF46<sub>1-433</sub>, ORF46<sub>433-608</sub>, ORF46, 961c etc.).

20 [0049] Particularly preferred are: (a) a hybrid protein comprising 919 and 287; (b) a hybrid protein comprising 953 and 287; (c) a hybrid protein comprising 287 and ORF46.1; (d) a hybrid protein comprising ORF1 and ORF46.1; (e) a hybrid protein comprising 919 and ORF46.1; (f) a hybrid protein comprising ORF46.1 and 919; (g) a hybrid protein comprising ORF46.1, 287 and 919; (h) a hybrid protein comprising 919 and 519; and (i) a hybrid protein comprising ORF97 and 225. Further embodiments are shown in Figure 14.

25 [0050] Where 287 is used, it is preferably at the C-terminal end of a hybrid; if it is to be used at the N-terminus, if is preferred to use a  $\Delta G$  form of 287 is used (e.g. as the N-terminus of a hybrid with ORF46.1, 919, 953 or 961).

30 [0051] Where 287 is used, this is preferably from strain 2996 or from strain 394/98.

[0052] Where 961 is used, this is preferably at the N-terminus. Domain forms of 961 may be used.

35 [0053] Alignments of polymorphic forms of ORF46, 287, 919 and 953 are disclosed in WO00/66741. Any of these polymorphs can be used according to the present invention.

### Temperature

40 [0054] In a sixth approach to heterologous expression, proteins of the invention are expressed at a low temperature.

[0055] Expressed Neisserial proteins (e.g. 919) may be toxic to *E.coli*, which can be avoided by expressing the toxic protein at a temperature at which its toxic activity is not manifested.

[0056] Thus the present invention provides a method for the heterologous expression of a protein of the invention, in which expression of a protein of the invention is carried out at a temperature at which a toxic activity of the protein is not manifested.

[0057] A preferred temperature is around 30°C. This is particularly suited to the expression of 919.

### Mutations

45 [0058] As discussed above, expressed Neisserial proteins may be toxic to *E coli*. This toxicity can be avoided by mutating the protein to reduce or eliminate the toxic activity. In particular, mutations to reduce or eliminate toxic enzymatic activity can be used, preferably using site-directed mutagenesis.

[0059] In a seventh approach to heterologous expression, therefore, an expressed protein is mutated to reduce or eliminate toxic activity.

[0060] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which protein is mutated to reduce or eliminate toxic activity.

50 [0061] The method is preferably used for the expression of protein 907, 919 or 922. A preferred mutation in 907 is at Glu-117 (e.g. Glu→Gly); preferred mutations in 919 are at Glu-255 (e.g. Glu→Gly) and/or Glu-323 (e.g. Glu→Gly); preferred mutations in 922 are at Glu-164 (e.g. Glu→Gly), Ser-213 (e.g. Ser→Gly) and/or Asn-348 (e.g. Asn→Gly).

### Alternative vectors

55 [0062] In a eighth approach to heterologous expression, an alternative vector used to express the protein. This may be to improve expression yields, for instance, or to utilise plasmids that are already approved for GMP use.

[0063] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which

an alternative vector is used. The alternative vector is preferably pSM214, with no fusion partners. Leader peptides may or may not be included.

[0064] This approach is particularly useful for protein 953. Expression and localisation of 953 with its native leader peptide expressed from pSM214 is much better than from the pET vector.

[0065] pSM214 may also be used with: ΔG287, Δ2-287, Δ3-287, Δ4-287, Orf46.1, 961L, 961, 96 1 (MC58), 96 1 c, 96 1 c-L, 919, 953 and ΔG287-Orf46.1.

[0066] Another suitable vector is pET-24b (Novagen; uses kanamycin resistance), again using no fusion partners. pET-24b is preferred for use with: ΔG287K, Δ2-287K, Δ3-287K, Δ4-287K,

[0067] Orf46.1-K, Orf46A-K, 961-K (MC58), 961a-K, 961b-K, 961c-K, 961c-L-K, 961d-K, ΔG287-919-K, ΔG287-Orf46.1-K and ΔG287-961-K.

#### ***Multimeric form***

[0068] In a ninth approach to heterologous expression, a protein is expressed or purified such that it adopts a particular multimeric form.

[0069] This approach is particularly suited to protein 953. Purification of one particular multimeric form of 953 (the monomeric form) gives a protein with greater bactericidal activity than other forms (the dimeric form).

[0070] Proteins 287 and 919 may be purified in dimeric forms.

[0071] Protein 961 may be purified in a 180kDa oligomeric form (e.g. a tetramer).

#### ***Lipidation***

[0072] In a tenth approach to heterologous expression, a protein is expressed as a lipidated protein.

[0073] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which the protein is expressed as a lipidated protein.

[0074] This is particularly useful for the expression of 919, 287, ORF4, 406, 576-1, and ORF25. Polymorphic forms of 919, 287 and ORF4 are disclosed in WO00/66741.

[0075] The method will typically involve the use of an appropriate leader peptide without using an N-terminal fusion partner.

#### ***C-terminal deletions***

[0076] In an eleventh approach to heterologous expression, the C-terminus of a protein of the invention is mutated. In addition, it is preferred that no fusion partner is used.

[0077] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) the protein's C-terminus region is mutated and, optionally, (b) no fusion partner is used.

[0078] The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; manipulating said nucleic acid to mutate nucleotides that encode the protein's C-terminus portion. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector. The first amino acid of the expressed protein will be that of the mature native protein.

[0079] The mutation may be a substitution, insertion or, preferably, a deletion.

[0080] This method can increase the levels of expression, particularly for proteins 730, ORF29 and ORF46. For protein 730, a C-terminus region of around 65 to around 214 amino acids may be deleted; for ORF46, the C-terminus region of around 175 amino acids may be deleted; for ORF29, the C-terminus may be deleted to leave around 230-370 N-terminal amino acids.

#### ***Leader peptide mutation***

[0081] In a twelfth approach to heterologous expression, the leader peptide of the protein is mutated. This is particularly useful for the expression of protein 919.

[0082] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which the protein's leader peptide is mutated.

[0083] The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; and manipulating said nucleic acid to mutate nucleotides within the leader peptide. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector.

**Poly-glycine deletion**

[0084] In a thirteenth approach to heterologous expression, poly-glycine stretches in wild-type sequences are mutated. This enhances protein expression.

5 [0085] The poly-glycine stretch has the sequence (Gly)<sub>n</sub>, where n≥4 (e.g. 5, 6, 7, 8, 9 or more). This stretch is mutated to disrupt or remove the (Gly)<sub>n</sub>. This may be by deletion (e.g. CGGGGS→CGGGS, CGGS, CGS or CS), by substitution (e.g. CGGGGS→CGXGGS, CGXXGS, CGXGXS etc.), and/or by insertion (e.g. CGGGGS→CGGXGGS, CGXGGGS, etc.).

10 [0086] This approach is not restricted to Neisserial proteins - it may be used for any protein (particularly bacterial proteins) to enhance heterologous expression. For Neisserial proteins, however, it is particularly suitable for expressing 287, 741, 983 and Tbp2. An alignment of polymorphic forms of 287 is disclosed in WO00/66741.

[0087] Thus the invention provides a method for the heterologous expression of a protein of the invention, in which (a) a poly-glycine stretch within the protein is mutated.

15 [0088] The method will typically involve the steps of: obtaining nucleic acid encoding a protein of the invention; and manipulating said nucleic acid to mutate nucleotides that encode a poly-glycine stretch within the protein sequence. The resulting nucleic acid may be inserted into an expression vector, or may already be part of an expression vector.

[0089] Conversely, the opposite approach (i.e. introduction of poly-glycine stretches) can be used to suppress or diminish expression of a given heterologous protein.

20 **Heterologous host**

[0090] Whilst expression of the proteins of the invention may take place in the native host (i.e. the organism in which the protein is expressed in nature), the present invention utilises a heterologous host. The heterologous host may be prokaryotic or eukaryotic. It is preferably *E.coli*, but other suitable hosts include *Bacillus subtilis*, *Vibrio cholerae*, *Salmonella typhi*, *Salmonella typhimurium*, *Neisseria meningitidis*, *Neisseria gonorrhoeae*, *Neisseria lactamica*, *Neisseria cinerea*, *Mycobacteria* (e.g. *M.tuberculosis*), yeast etc.

**Vectors etc.**

30 [0091] As well as the methods described above, the invention provides (a) nucleic acid and vectors useful in these methods (b) host cells containing said vectors (c) proteins expressed or expressable by the methods (d) compositions comprising these proteins, which may be suitable as vaccines, for instance, or as diagnostic reagents, or as immunogenic compositions (e) these compositions for use as medicaments (e.g. as vaccines) or as diagnostic reagents (f) the use of these compositions in the manufacture of (1) a medicament for treating or preventing infection due to Neisserial bacteria 35 (2) a diagnostic reagent for detecting the presence of Neisserial bacteria or of antibodies raised against Neisserial bacteria, and/or (3) a reagent which can raise antibodies against Neisserial bacteria and (g) a method of treating a patient, comprising administering to the patient a therapeutically effective amount of these compositions.

**Sequences**

40 [0092] The invention also provides a protein or a nucleic acid having any of the sequences set out in the following examples. It also provides proteins and nucleic acid having sequence identity to these. As described above, the degree of 'sequence identity' is preferably greater than 50% (eg. 60%, 70%, 80%, 90%, 95%, 99% or more).

[0093] Furthermore, the invention provides nucleic acid which can hybridise to the nucleic acid disclosed in the examples, preferably under "high stringency" conditions (eg. 65°C in a 0.1xSSC, 0.5% SDS solution).

[0094] The invention also provides nucleic acid encoding proteins according to the invention.

[0095] It should also be appreciated that the invention provides nucleic acid comprising sequences complementary to those described above (eg. for antisense or probing purposes).

[0096] Nucleic acid according to the invention can, of course, be prepared in many ways (eg. by chemical synthesis, from genomic or cDNA libraries, from the organism itself etc.) and can take various forms (eg. single stranded, double stranded, vectors, probes etc.).

[0097] In addition, the term "nucleic acid" includes DNA and RNA, and also their analogues, such as those containing modified backbones, and also peptide nucleic acids (PNA) etc.

55 **BRIEF DESCRIPTION OF DRAWINGS**

[0098]

Figures 1 and 2 show constructs used to express proteins using heterologous leader peptides.

Figure 3 shows expression data for ORF1, and Figure 4 shows similar data for protein 961.

5 Figure 5 shows domains of protein 287, and Figures 6 & 7 show deletions within domain A.

Figure 8 shows domains of protein 564.

10 Figure 9 shows the *PhoC* reporter gene driven by the 919 leader peptide, and Figure 10 shows the results obtained using mutants of the leader peptide.

Figure 11 shows insertion mutants of protein 730 (A: 730-C1; B: 730-C2).

15 Figure 12 shows domains of protein 961.

Figure 13 shows SDS-PAGE of ΔG proteins. Dots show the main recombinant product.

Figure 14 shows 26 hybrid proteins according to the invention.

## 20 MODES FOR CARRYING OUT THE INVENTION

### *Example 1- 919 and its leader peptide*

[0099] Protein 919 from *N.meningitidis* (serogroup B, strain 2996) has the following sequence:

1	MKKYLFR <b>AAL</b> YGIAAA <b>I</b> LAA CQS <span style="font-family: monospace;">KSIQTFP QPDTSVINGP DRPGVIPDPA</span>
51	GTTVGGGGAV YTVP <i>HLSLP</i> HWAAQDFAKS LQSFRLG <b>CAN</b> LKNRQGWQDV
101	CAQAFQTPVH SFQAKQFFER YFTP <i>WQVAGN</i> GSLAGTVTGY YEPVLKGDDR
151	RTAQARFPIY GIPDDFISVP LPAGLRS <b>GKA</b> LVRIRQTGKN SGTIDNTGGT
201	HTADLSRFPI TARTTAIKGR FEGRFLPYH TRNQINGGAL DGKAPILGYA
251	EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL
301	KLGQTSMQGI KAYMRQNPNR LAEVLGQNPS YIFFRELAGS SNDGPVGALG
351	TPLMGEYAGA VDRHYITLGA PLFVATAH <i>PV</i> TRKALNRLIM AQDTGSAIKG
401	AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P*

35 [0100] The leader peptide is underlined.

[0101] The sequences of 919 from other strains can be found in Figures 7 and 18 of WO00/66741.

[0102] Example 2 of WO99/57280 discloses the expression of protein 919 as a His-fusion in *E.coli*.

[0103] The protein is a good surface-exposed immunogen.

40 [0104] Three alternative expression strategies were used for 919:

- 1) 919 without its leader peptide (and without the mature N-terminal cysteine) and without any fusion partner ('919<sup>untagged</sup>');

1	QSKSIQTFP QPDTSVINGP DRPGVIPDPA GTTVGGGGAV YTVP <i>HLSLP</i>
50	HWAAQDFAKS LQSFRLG <b>CAN</b> LKNRQGWQDV CAQAFQTPVH SFQAKQFFER
100	YFTP <i>WQVAGN</i> GSLAGTVTGY YEPVLKGDDR RTAQARFPIY GIPDDFISVP
150	LPAGLRS <b>GKA</b> LVRIRQTGKN SGTIDNTGGT HTADLSRFPI TARTTAIKGR
200	FEGRFLPYH TRNQINGGAL DGKAPILGYA EDPVELFFMH IQGSGRLKTP
250	SGKYIRIGYA DKNEHPYVSI GRYMADKGYL KLGQTSMQGI KAYMRQNPNR
300	LAEVLGQNPS YIFFRELAGS SNDGPVGALG TPLMGEYAGA VDRHYITLGA
350	PLFVATAH <i>PV</i> TRKALNRLIM AQDTGSAIKG AVRVDYFWGY GDEAGELAGK
400	QKTTGYVWQL LPNGMKPEYR P*

55 The leader peptide and cysteine were omitted by designing the 5'-end amplification primer downstream from the predicted leader sequence.

2) 919 with its own leader peptide but without any fusion partner ('919L');

3) 919 with the leader peptide (MKTFFKTL*SAAALALI*LAA) from ORF4 ('919LOrf4').

1    MKTFFKTLS AAALALILAA CQSJKSIQTFP QPDTSVINGP DRPVGIPDPA  
 5    GTTVGGGGAV YTVVPHSLP HWAAQDFAKS LQSFRILGCAN LKNRQGWQDV  
 10   CAQAFQTPVH SFQAKQFFER YFTPWQVAGN GSLAGTVTGY YEPVLKGDDR  
 15   RTAQARFPPI GIPDDFISVP LPAGLRSGKA LVRIRQTGKN SGTIDNTGGT  
 20   HTADLSRFPI TARTTAIKGR FEGSRFLPYH TRNQINGGAL DGKAPILGYA  
 25   EDPVELFFMH IQGSGRLKTP SGKYIRIGYA DKNEHPYVSI GRYMADKGYL  
 30   KLGQTSQMGI KSYMQRNPQR LAEVLGQNPS YIFFRELAGS SNDGPVGALG

10

350   TPLMGEYAGA VDRHYITLGA PLFVATAHPV TRKALNRLIM AQDTGSAIKG  
 400   AVRVDYFWGY GDEAGELAGK QKTTGYVWQL LPNGMKPEYR P\*

15   To make this construct, the entire sequence encoding the ORF4 leader peptide was included in the 5'-primer as a tail (primer 919Lorf4 For). A *N*hel restriction site was generated by a double nucleotide change in the sequence coding for the ORF4 leader (no amino acid changes), to allow different genes to be fused to the ORF4 leader peptide sequence. A stop codon was included in all the 3'-end primer sequences.

- 20   [0105] All three forms of the protein were expressed and could be purified.  
 [0106] The '919L' and '919LOrf4' expression products were both lipidated, as shown by the incorporation of [<sup>3</sup>H]-palmitate label. 919untagged did not incorporate the <sup>3</sup>H label and was located intracellularly.  
 [0107] 919LOrf4 could be purified more easily than 919L. It was purified and used to immunise mice. The resulting sera gave excellent results in FACS and ELISA tests, and also in the bactericidal assay. The lipoprotein was shown to be localised in the outer membrane.  
 25   [0108] 919untagged gave excellent ELISA titres and high serum bactericidal activity. FACS confirmed its cell surface location.

#### **Example 2 — 919 and expression temperature**

30   [0109] Growth of *E.coli* expressing the 919LOrf4 protein at 37°C resulted in lysis of the bacteria. In order to overcome this problem, the recombinant bacteria were grown at 30°C. Lysis was prevented without preventing expression.

#### **Example 3 - mutation of 907, 919 and 922**

- 35   [0110] It was hypothesised that proteins 907, 919 and 922 are murein hydrolases, and more particularly lytic transglycosylases. Murein hydrolases are located on the outer membrane and participate in the degradation of peptidoglycan.  
 [0111] The purified proteins 919untagged, 919Lorf4, 919-His (*i.e.* with a C-terminus His-tag) and 922-His were thus tested for murein hydrolase activity [Ursinus & Holtje (1994) J.Bact. 176:338-343]. Two different assays were used, one determining the degradation of insoluble murein sacculus into soluble muropeptides and the other measuring breakdown of poly(MurNAc-GlcNAc)<sub>n>30</sub> glycan strands.  
 40   [0112] The first assay uses murein sacci radiolabelled with meso-2,6-diamino-3,4,5-[<sup>3</sup>H]pimelic acid as substrate. Enzyme (3-10 µg total) was incubated for 45 minutes at 37°C in a total volume of 100µl comprising 10mM Tris-maleate (pH 5.5), 10mM MgCl<sub>2</sub>, 0.2% v/v Triton X-100 and [<sup>3</sup>H]A<sub>2</sub>pm labelled murein sacci (about 10000cpm). The assay mixture was placed on ice for 15 minutes with 100 µl of 1% w/v N-acetyl-N,N,N-trimethylammonium for 15 minutes and precipitated material pelleted by centrifugation at 10000g for 15 minutes. The radioactivity in the supernatant was measured by liquid scintillation counting. *E.coli* soluble lytic transglycosylase Slt70 was used as a positive control for the assay; the negative control comprised the above assay solution without enzyme.  
 45   [0113] All proteins except 919-His gave positive results in the first assay.  
 [0114] The second assay monitors the hydrolysis of poly(MurNAc-GlcNAc)glycan strands. Purified strands, poly(MurNAc-GlcNAc)<sub>n>30</sub> labelled with N-acetyl-D-1-[<sup>3</sup>H]glucosamine were incubated with 3µg of 919L in 10 mM Tris-maleate (pH 5.5), 10 mM MgCl<sub>2</sub> and 0.2% v/v Triton X-100 for 30 min at 37°C. The reaction was stopped by boiling for 5 minutes and the pH of the sample adjusted to about 3.5 by addition of 10µl of 20% v/v phosphoric acid. Substrate and product were separated by reversed phase HPLC on a Nucleosil 300 C<sub>18</sub> column as described by Harz et. al. [Anal. Biochem. (1990) 190:120-128]. The *E.coli* lytic transglycosylase Mlt A was used as a positive control in the assay. The negative control was performed in the absence of enzyme.  
 55   [0115] By this assay, the ability of 919LOrf4 to hydrolyse isolated glycan strands was demonstrated when anhydrodisaccharide subunits were separated from the oligosaccharide by HPLC.

**[0116]** Protein 919Lorf4 was chosen for kinetic analyses. The activity of 919Lorf4 was enhanced 3.7-fold by the addition of 0.2% v/v Triton X-100 in the assay buffer. The presence of Triton X-100 had no effect on the activity of 919untagged. The effect of pH on enzyme activity was determined in Tris-Maleate buffer over a range of 5.0 to 8.0. The optimal pH for the reaction was determined to be 5.5. Over the temperature range 18°C to 42°C, maximum activity was observed at 37°C. The effect of various ions on murein hydrolase activity was determined by performing the reaction in the presence of a variety of ions at a final concentration of 10mM. Maximum activity was found with Mg<sup>2+</sup>, which stimulated activity 2.1-fold. Mn<sup>2+</sup> and Ca<sup>2+</sup> also stimulated enzyme activity to a similar extent while the addition Ni<sup>2+</sup> and EDTA had no significant effect. In contrast, both Fe<sup>2+</sup>and Zn<sup>2+</sup> significantly inhibited enzyme activity.

[0117] The structures of the reaction products resulting from the digestion of unlabelled *E.coli* murein sacculus were analysed by reversed-phase HPLC as described by Glauner [Anal. Biochem. (1988) 172:451-464]. Murein sacculi digested with the muramidase Cellosyl were used to calibrate and standardise the Hypersil ODS column. The major reaction products were 1,6 anhydrodisaccharide tetra and tri peptides, demonstrating the formation of 1,6 anhydromuramic acid intramolecular bond.

[0118] These results demonstrate experimentally that 919 is a murein hydrolase and in particular a member of the lytic transglycosylase family of enzymes. Furthermore the ability of 922-His to hydrolyse murein sacculi suggests this protein is also a lytic transglycosylase.

**[0119]** This activity may help to explain the toxic effects of 919 when expressed in *E.coli*.

**[0119]** This activity may help to explain the toxic effects of S-15 when expressed in E. coli.

**[0120]** In order to eliminate the enzymatic activity, rational mutagenesis was used. 907, 919 and 922 show fairly low homology to three membrane-bound lipidated murein lytic transglycosylases from *E.coli*.

**919** (441aa) is 27.3% identical over 440aa overlap to *E.coli* MLTA (P46885).

**922** (369aa) is 38.7% identical over 310aa overlap to *E.coli* MLTB (P41052); and

907-2 (207aa) is 26.8% identical over 149aa overlap to *E.coli* MLTC (P52066).

25 907-2 also shares homology with *E.coli* MLTD (P23931) and SlT70 (P03810), a soluble lytic transglycosylase that is located in the periplasmic space. No significant sequence homology can be detected among 919, 922 and 907-2, and the same is true among the corresponding MLTA, MLTB and MLTC proteins.

[0121] Crystal structures are available for Slt70 [1QTEA; 1QTEB; Thunnissen et al. (1995) Biochemistry 34: 12729-12737] and for Slt35 [1LTM; 1QUS; 1QUT; van Asselt et al. (1999) Structure Fold Des 7:1167-80] which is a soluble form of the 40kDa MLTB.

[0122] The catalytic residue (a glutamic acid) has been identified for both Slt70 and MLTB.

**[0123]** In the case of Slt70, mutagenesis studies have demonstrated that even a conservative substitution of the catalytic Glu505 with a glutamine (Gln) causes the complete loss of enzymatic activity. Although Slt35 has no obvious sequence similarity to Slt70, their catalytic domains shows a surprising similarity. The corresponding catalytic residue in MLTB is Glu162.

**[0124]** Another residue which is believed to play an important role in the correct folding of the enzymatic cleft is a well-conserved glycine (Gly) downstream of the glutamic acid. Recently, Terrak et al. [Mol. Microbiol. (1999) 34:350-64] have suggested the presence of another important residue which is an aromatic amino acid located around 70-75 residues downstream of the catalytic glutamic acid.

[0125] Sequence alignment of Slt70 with 907-2 and of MLTB with 922 were performed in order to identify the corresponding catalytic residues in the MenB antigens.

[0126] The two alignments in the region of the catalytic domain are reported below:

45 907-2/Slt70:

	90	100	110	▼120	130	140
907-2.pep	ERRRLLVNIQYESRAG--LDTQIVLGLIEVESAFRQYAISGV <u>GARG</u> GLMQVMPFWKNYIG					
slty_ecoli	: : : : : : : :    :             :   :					
	ERFPLAYNDLFKRYTSGKEIPQSYAMAIARQESAWNPVKSPVGASGLMQIMP GTATHTV					
	480	490	500	▲ 510	520	530
				<b>GLU505</b>		

922/MLTB

5 922.pep  mltb_ecoli	150        160     ▼    170        180        190        200 VAQKYGVPAELIVAVIGIETNY <b>G</b> KNT <b>G</b> SFRVADALATLGFDYPRRAGFFQKELVELLKLA :         :: :: ::   :    :    :: :: : :   :    : AWQVYGVPPPEIIVGIIGVETRWGRVMGKTRILDA(LS)FNYPRRAEYFSGELET(FLLMA 150        160     ▲    170        180        190        200 <b>GLU162</b>
10 922.pep  mltb_ecoli	210        220        230        240        250        260 KEEGGDVFAFKGSYAGAMGMPQFMPSS <b>Y</b> RKAWAVDYDGDGH(RDIWGNVGDVAA(S)VANYMKQ ::    :   :  :: :   :: :    :: : :   ::     :    :    :: : RDEQDDPLNLKGSFAGAMGYGQFMPSS <b>Y</b> KQYAVDFSGDGHINLWDPV-DAIGSVANYFKA 210        220        230        240        250        260

- 15 [0127] From these alignments, it results that the corresponding catalytic glutamate in 907-2 is Glu117, whereas in 922 is Glu164. Both antigens also share downstream glycines that could have a structural role in the folding of the enzymatic cleft (in bold), and 922 has a conserved aromatic residue around 70aa downstream (in bold).
- [0128] In the case of protein 919, no 3D structure is available for its *E.coli* homologue MLTA, and nothing is known about a possible catalytic residue. Nevertheless, three amino acids in 919 are predicted as catalytic residues by alignment with MLTA:

919/MLTA

25 919.pep  mlta_ecoli.p	240        250     ▼    260     □    270     □    280        290 ALDGKAPILGYAEDPVELFFMHIQGSGLRKTPSGKYIRI-GYADKNEHPYVSIGRYMADK   :      : :: :   :    ::   :              :    : ALSDKY-ILAYSNSIMDNFIMDVQGSGYIDFGDGSPLNFSYAGKNGHAYRSIGKVLI(DR 170        180        190        200        210
30  35 919.pep  mlta_ecoli.p	300        310        320     ▼    330□ □□    340        350        ◊ GYLKLGQTSMQGIKS(Y)MRQNPO-RLAEVLGQNP(S)YIFFREL(G)SSNDGPV-GALGTPLMG    :   :    ! :   :   :   :    :        :    :            :    : GEVKKEDMSMQAIRHWGETHSEAEVRELLEQNPSFVFFKPKQSFA---PVKGASAVPLVG 220        230        240        250        260        270
40 919.pep  mlta_ecoli.p	360     ▼    ◊    380        390        400        ◊◊410 EYAGAVDRHYITLGAPLFVATAHPVTRKALN----RLIMAQDTGSAIKGAVRVDYFWGY :            :   :: :    :            :            :    : RASVASDRSIIPPGTTLAEVPLLDNNGKFNGQYELRLMVALDVGGAIKGQ-HFDIYQGI 280        290        300        310        320        330
45 919.pep  mlta_ecoli.p	420        ◊ GDEAGELAGKQKTTGYVWQLP        :       :        : GPEAGHRAGWYNHYGRVWVLKT 340        350

- 50 [0129] The three possible catalytic residues are shown by the symbol ▼:

- 1) Glu255 (Asp in MLTA), followed by three conserved glycines (Gly263, Gly265 and Gly272) and three conserved aromatic residues located approximately 75-77 residues downstream. These downstream residues are shown by ◊.
- 2) Glu323 (conserved in MLTA), followed by 2 conserved glycines (Gly347 and Gly355) and two conserved aromatic residues located 84-85 residues downstream (Tyr406 or Phe407). These downstream residues are shown by ◊◊.
- 3) Asp362 (instead of the expected Glu), followed by one glycine (Gly 369) and a conserved aromatic residue

(Trp428). These downstream residues are shown by ○.

[0130] Alignments of polymorphic forms of 919 are disclosed in WO00/66741.

[0131] Based on the prediction of catalytic residues, three mutants of the 919 and one mutant of 907, containing each a single amino acid substitution, have been generated. The glutamic acids in position 255 and 323 and the aspartic acids in position 362 of the 919 protein and the glutamic acid in position 117 of the 907 protein, were replaced with glycine residues using PCR-based SDM. To do this, internal primers containing a codon change from Glu or Asp to Gly were designed:

10	Primers	Sequences	Codon change
919-E255 for 919-E255 rev	CGAAGACCCCGTC <u>Ggt</u> CTTTTTTTATG GTGCATAAAAAAAAGacCGACGGGGTCT	GAA → Ggt	
919-E323 for 919-E323 rev	AACGCCTCGCC <u>Ggt</u> GTTTGCGGTCA TTTGACCCAAAACacCGGCGAGGCG	GAA → Ggt	
919-D362 for 919-D362 rev	TGCCGGCGCAGTC <u>Ggt</u> CGGCACTACA TAATGTAGTGCCGacCGACTGCGCCG	GAC → Ggt	
907-E117 for 907-E117 rev	TGATTGAGGTG <u>Ggt</u> AGCGCGTCCG GGCGGAACGCGCTacCCACCTCAAT	GAA → Ggt	
Underlined nucleotides code for glycine; the mutated nucleotides are in lower case.			

[0132] To generate the 919-E255, 919-E323 and 919-E362 mutants, PCR was performed using 20ng of the pET 919-LOrf4 DNA as template, and the following primer pairs:

- 1) Orf4L for / 919-E255 rev
- 2) 919-E255 for / 919L rev
- 3) Orf4L for / 919-E323 rev
- 4) 919-E323 for / 919L rev
- 5) Orf4L for / 919-D362 rev
- 6) 919-D362 for / 919L rev

The second round of PCR was performed using the product of PCR 1-2, 3-4 or 5-6 as template, and as forward and reverse primers the "Orf4L for" and "919L rev" respectively.

For the mutant 907-E117, PCR have been performed using 200ng of chromosomal DNA of the 2996 strain as template and the following primer pairs:

- 7) 907L for / 907-E117 rev
- 8) 907-E 117 for / 907L rev

[0133] The second round of PCR was performed using the products of PCR 7 and 8 as templates and the oligos "907L for" and "907L rev" as primers.

[0134] The PCR fragments containing each mutation were processed following the standard procedure, digested with *Nde*I and *Xho*I restriction enzymes and cloned into pET-21b+ vector. The presence of each mutation was confirmed by sequence analysis.

[0135] Mutation of Glu117 to Gly in 907 is carried out similarly, as is mutation of residues Glu164, Ser213 and Asn348 in 922.

[0136] The E255G mutant of 919 shows a 50% reduction in activity; the E323G mutant shows a 70% reduction in activity; the E362G mutant shows no reduction in activity.

#### 50 Example 4 - multimeric form

[0137] 287-GST, 919<sup>untagged</sup> and 953-His were subjected to gel filtration for analysis of quaternary structure or preparative purposes. The molecular weight of the native proteins was estimated using either FPLC Superose 12 (H/R 10/30) or Superdex 75 gel filtration columns (Pharmacia). The buffers used for chromatography for 287, 919 and 953 were 50 mM Tris-HCl (pH 8.0), 20 mM Bicine (pH 8.5) and 50 mM Bicine (pH 8.0), respectively.

[0138] Additionally each buffer contained 150-200 mM NaCl and 10% v/v glycerol. Proteins were dialysed against the appropriate buffer and applied in a volume of 200μl. Gel filtration was performed with a flow rate of 0.5 - 2.0 ml/min and

the eluate monitored at 280nm. Fractions were collected and analysed by SDS-PAGE. Blue dextran 2000 and the molecular weight standards ribonuclease A, chymotrypsin A ovalbumin, albumin (Pharmacia) were used to calibrate the column. The molecular weight of the sample was estimated from a calibration curve of  $K_{av}$  vs. log  $M_r$  of the standards. Before gel filtration, 287-GST was digested with thrombin to cleave the GST moiety.

[0139] The estimated molecular weights for 287, 919 and 953-His were 73 kDa, 47 kDa and 43 kDa respectively. These results suggest 919 is monomeric while both 287 and 953 are principally dimeric in their nature. In the case of 953-His, two peaks were observed during gel filtration. The major peak (80%) represented a dimeric conformation of 953 while the minor peak (20%) had the expected size of a monomer. The monomeric form of 953 was found to have greater bactericidal activity than the dimer.

#### 10 Example 5 - pSM214 and pET-24b vectors

[0140] 953 protein with its native leader peptide and no fusion partners was expressed from the pET vector and also from pSM214 [Velati Bellini et al. (1991) J. Biotechnol. 18, 177-192].

[0141] The 953 sequence was cloned as a full-length gene into pSM214 using the *E. coli* MM294-1 strain as a host. To do this, the entire DNA sequence of the 953 gene (from ATG to the STOP codon) was amplified by PCR using the following primers:

953L for/2 CCGGAATTCTTATGAAAAAAATCATCTTCGCCGC Eco RI

953L rev/2 GCCCAAGCTTTATTGGCTGCCTCGATT Hind III

which contain EcoRI and HindIII restriction sites, respectively. The amplified fragment was digested with EcoRI and HindIII and ligated with the pSM214 vector digested with the same two enzymes. The ligated plasmid was transformed into *E.coli* MM294-1 cells (by incubation in ice for 65 minutes at 37° C) and bacterial cells plated on LB agar containing 20µg/ml of chloramphenicol.

[0142] Recombinant colonies were grown over-night at 37°C in 4 ml of LB broth containing 20 µg/ml of chloramphenicol; bacterial cells were centrifuged and plasmid DNA extracted as and analysed by restriction with EcoRI and HindIII. To analyse the ability of the recombinant colonies to express the protein, they were inoculated in LB broth containing 20µg/ml of chloramphenicol and let to grow for 16 hours at 37°C. Bacterial cells were centrifuged and resuspended in PBS. Expression of the protein was analysed by SDS-PAGE and Coomassie Blue staining.

[0143] Expression levels were unexpectedly high from the pSM214 plasmid.

[0144] Oligos used to clone sequences into pSM-214 vectors were as follows:

35	<b>ΔG287</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTTATG-TCGCCCGATGTTAAATCGCGGA	EcoRI
		Rev	GCCCAAGCTT-TCAATCCTGCTTTTTGCCG	HindIII
40	<b>Δ2 287</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTTATG-AGCCAAGATATGGCGGCAGT	EcoRI
		Rev	GCCCAAGCTT-TCAATCCTGCTTTTTGCCG	HindIII
45	<b>Δ3 287</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTTATG-TCCGCCAATCCGCAAATCA	EcoRI
		Rev	GCCCAAGCTT-TCAATCCTGCTTTTTGCCG	HindIII
50	<b>Δ4 287</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTTATG-GGAAGGGTTGATTGGCTAATG	EcoRI
		Rev	GCCCAAGCTT-TCAATCCTGCTTTTTGCCG	HindIII
55	<b>Orf46.1</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTTATG-TCAGATTGGCAAACGATTCTT	EcoRI
		Rev	GCCCAAGCTT-TTACGTATCATATTCACGTGCTTC	HindIII
60	<b>ΔG287-Orf46.1</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTTATG-TCGCCCGATGTTAAATCGCGGA	EcoRI
		Rev	GCCCAAGCTT-TTACGTATCATATTCACGTGCTTC	HindIII
65	<b>919</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTTATG-CAAAGCAAGAGCATCCAAACCT	EcoRI
		Rev	GCCCAAGCTT-TTACGGCGGTATCGGGCT	HindIII
70	<b>961L</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTATG-AAACACTTCCATCC	EcoRI
		Rev	GCCCAAGCTT-TTACCACTCGTAATTGAC	HindIII
75	<b>961</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTATG-GCCACAAGCGACGAC	EcoRI
		Rev	GCCCAAGCTT-TTACCACTCGTAATTGAC	HindIII

(continued)

5	<b>961c L</b> <b>pSM-214</b>	Fwd	CCGGAATTCTTATG-AAACACTTCCATCC	EcoRI
		Rev	GCCCAAGCTT-TCAACCCACGTTGAAGGTTG	HindIII
10	<b>961c</b> <b>pSM-214</b>	Fwd	CCGGAATTCTTATG-GCCACAAACGACGACG	EcoRI
		Rev	GCCCAAGCTT-TCAACCCACGTTGAAGGTTG	HindIII
15	<b>953</b> <b>(pSM-214)</b>	Fwd	CCGGAATTCTTATG-GCCACCTACAAAGTGGACGA	EcoRI
		Rev	GCCCAAGCTT-TTATTGTTGGCTGCCTCGATT	HindIII

These sequences were manipulated, cloned and expressed as described for 953L.

[0145] For the pET-24 vector, sequences were cloned and the proteins expressed in pET-24 as described below for pET21. pET2 has the same sequence as pET-21, but with the kanamycin resistance cassette instead of ampicillin cassette.

[0146] Oligonucleotides used to clone sequences into pET-24b vector were:

20	<b>ΔG 287 K</b>	Fwd	CGCGGATCCGCTAGC-CCCGATGTTAACCGGC §	Nhel
		Rev	CCCGCTCGAG-TCAATCCTGCTCTTTTGCC *	Xhol
25	<b>Δ2 287 K</b>	Fwd	CGCGGATCCGCTAGC-CAAGATATGGCGGGCAGT §	Nhel
		Rev	CGCGGATCCGCTAGC-GCCGAATCCGCAAATCA §	Nhel
30	<b>Δ4 287 K</b>	Fwd	CGCGCTAGC-GGAAGGGTTGATTGGCTAATGG §	Nhel
		Rev	GGGAATTCCATATG-GGCATTTCCCGCAAAATATC	Ndel
35	<b>Orf46.1 K</b>	Fwd	CCCGCTCGAG-TTACGTATCATATTCACGTGC	Xhol
		Rev	GGGAATTCCATATG-GGCATTTCCCGCAAAATATC	Ndel
40	<b>Orf46A K</b>	Fwd	CCCGCTCGAG-TTACGTATCATATTCACGTGC	Xhol
		Rev	GGGAATTCCATATG-GGCATTTCCCGCAAAATATC	Ndel
45	<b>961 K</b> <b>(MC58)</b>	Fwd	CGCGGATCCCATATG-GCCACAAACGACGA	Ndel
		Rev	CCCGCTCGAG-TTACCACTCGTAATTGAC	Xhol
50	<b>961a K</b>	Fwd	CGCGGATCCCATATG-GCCACAAACGACG	Ndel
		Rev	CCCGCTCGAG-TCATTTAGCAATATTATCTTTGTT	Xhol
55	<b>961b K</b>	Fwd	CGCGGATCCCATATG-AAAGCAAACAGTGCCGAC	Ndel
		Rev	CCCGCTCGAG-TTACCACTCGTAATTGAC	Xhol
	<b>961c K</b>	Fwd	CGCGGATCCCATATG-GCCACAAACGACG	Ndel
		Rev	CCCGCTCGAG-TTAACCCACGTTGAAGGT	Xhol
	<b>961cL K</b>	Fwd	CGCGGATCCCATATG-ATGAAACACTTCCATCC	Ndel
		Rev	CCCGCTCGAG-TTAACCCACGTTGAAGGT	Xhol
	<b>961d K</b>	Fwd	CGCGGATCCCATATG-GCCACAAACGACG	Ndel
		Rev	CCCGCTCGAG-TCAGTCTGACACTGTTTATCC	Xhol
	<b>ΔG 287-</b> <b>919 K</b>	Fwd	CCCGCTCGAG-TTACGGGCGGTATTGG	Xhol
		Rev	CGCGGATCCGCTAGC-CCCGATGTTAACCGGC	Nhel
	<b>ΔG 287-</b> <b>Orf46.1 K</b>	Fwd	CCCGCTCGAG-TTACGTATCATATTCACGTGC	Xhol
		Rev	CGCGGATCCGCTAGC-CCCGATGTTAACCGGC	Nhel
	<b>ΔG 287-</b>	Fwd	CCCGCTCGAG-TTACGTATCATATTCACGTGC	Xhol
		Rev	CGCGGATCCGCTAGC-CCCGATGTTAACCGGC	Nhel

(continued)

5	<b>961 K</b>	Rev	<u>CCCGCTCGAG-TTACCACTCGTAATTGAC</u>	Xhol
* This primer was used as a Reverse primer for all the 287 forms.				
§ Forward primers used in combination with the ΔG278 K reverse primer.				

**Example 6 - ORF1 and its leader peptide**

- 10 [0147] ORF1 from *N.meningitidis* (serogroup B, strain MC58) is predicted to be an outer membrane or secreted protein. It has the following sequence:

15	1 <u>MKTTDKRTTE</u> <u>THRKAPKTGR</u> <u>IRFSPAYLAI</u> <u>CLSGFILPOA</u> WAGHTYFGIN 51    YQYYRDFAEN    KGKFAVGAKD    IEVYNKKGEL    VKKSMTKAPM    IDFSVVSRNG 101    VAALVGDQYI    VSVAHNGGYN    NVDFGAEGRN    PDQHRTFYKI    VKRNNYKAGT 151    KGHPYGGDYH    MPRLHKFVTD    AEPVEMTSYM    DGRKYIDQNN    YPDRVRIGAG 201    RQYWRSDEDE    PNNRESSYHI    ASAYSWLVGG    NTFAQNGSGG    GTVNLGSEKI 251    KHSPYGFPLT    GGSFGDGSGP    MFIYDAQKQK    WLINGVLQTG    NPYIGKSNGF 301    QLVRKDWFYD    EIFAGDTHSV    FYEPRQNKGK    SFNDDNNGTG    KINAKHEHNS 351    LPNRLKTRTV    QLFNVSLSET    AREPVYHAAG    GVNSYRPRLN    NGENISFIDE 401    GKGEELITSN    INQGAGGLYF    QGDFTVSPEN    NETWQGAGVH    ISEDSTVTWK 451    VNGVANDRLS    KIGKGTLHVQ    AKGENQGSIS    VGDGTVIDQ    QADDKGKKQA 501    FSEIGLVSQR    GTVQLNADNQ    FNPDKLYFGF    RGGRLLDNLH    SLSFHRIQNT 551    DEGAMIVNHN    QDKESTVTIT    GNKDIATTGN    NNSLDSKKEI    AYNGWFGEKD 601    TTKTNGRLNL    VYQPAAEEDRT    LLLSGGTNLN    GNITQTNGKL    FFSGRPTPH 651    YNHLDNDHSQ    KEGIPRGEIV    WDNDWINRTF    KAENFQIKGG    QAVVSRNVAK 701    VKGDWHLNSNH    AQAVFGVAPH    QSHTICTRSW    WTGLTNCVEK    TITDDKVIA 751    LTKTDISGNV    DLADAHHLNL    TGLATLNGNL    SANGDTRYTV    SHNATQNGNL 801    SLVGNAQATF    NQATLNGNTS    ASGNASFNLS    DHAVQNGSLT    LSGNAKANVS 851    HSALNGNVSL    ADKAVFHES    SRFTGQISGG    KDTALHLKDS    EWTLPSGTEL 901    GNLNLDNATI    TLNSAYRHDA    AGAQTSATD    APRRRSRRSR    RSLLSVTPPT 951    SVESRFNTLT    VNGKLNGQGT    FRFRMSELFY    RSDKLKLAES    SEGYTTLAVN 1001    NTGNEPASLE    QLTVEVGKDN    KPLSENLNFT    LQNEHVDAGA    WRYQLIRKD 1051    EFRLHNPVKE    QELSDKLGKA    EAKKQAEKDN    AQSLDALIAA    GRDAVEKTES 1101    VAEPARQAGG    ENVGIMQAAE    EKKRVQADKD    TALAKQREAE    TRPATTAFPR 1151    ARRARRDLQP    LQPQPQPQPQ    RDLISRYANS    GLSEFSATLN    SVFAVQDELD 1201    RVFAEDRRNA    VWTSGIRDTK    HYRSQDFRAY    RQQTDLRQIG    MQKNLGSGRV 1251    GILFSHNRTE    NTFDDGIGNS    ARLAHGAVFG    QYGIDRFYIG    ISAGAGFSSG 1301    SLSDGIGGKI    RRRVILHYGIQ    ARYRAGFGGF    GIEPHIGATR    YFVQKADYRY 1351    ENVNIATPGL    AFNRYRAGIK    ADYSFKPAQH    ISITPYLSLS    YTDAASGKVR 1401    TRVNTAVLAQ    DFGKTRSAEW    GVNAEIKGFT    LSLHAAAAGK    PQLEAQHSAG 1451    IKLGYRW*
----	---

40 The leader peptide is underlined.

[0148] A polymorphic form of ORF1 is disclosed in WO99/55873.

[0149] Three expression strategies have been used for ORF1:

- 45 1) ORF1 using a His tag, following WO99/24578 (ORF1-His);  
2) ORF1 with its own leader peptide but without any fusion partner ('ORF1L'); and  
3) ORF1 with the leader peptide (MKKTAIAIAVALAGFATVAQAA) from *E.coli* OmpA ('Orf1LOmpA');

50

55

5 MKKTAIAIAVALAGFATVAQQAASAGHTYFGINYQYYRDAENKGKFAVGAKDIEVYNKKGELVGKSMTKAPMIDFSV  
 VSRNGVAALVGDOIIVSVAHNGGYNNVDFGAERNRDQHRTYKIVKRNNYKAGTKGHPYGGDYHMPRLHKFVTDAE  
 PVEMTSYMDGRKYIDQNYPDRVRIGAGRQYWRSDEDE PNNRESSYHIASAYSWLVGGNTFAQNGSGGTNVNLGSEK  
 IKHSPYGFLPTGGSFGDSGSPMFYDAQKQKWINGVLQTNPYIGKSNGFQLVRKDWFYDEI FAGDTHSVFYEPHQ  
 NGKYSFNDDNNNTGKINAKHEHNSLPNRLKRTVQLFNVSLSETAREPVYHAAGGVNSYRPRLNNGENISFIDEKGK  
 ELILTSNINQGAGGLYFQGDFTVSPENNWTQGAGVHISEDSTVTWKVNGVANDRLSKIGKGTLLHVQAKGENQGSIS  
 VGDGTVIDQQADDKGKQAFSEIGLVSGRGTVQLNADNFNQFPDKLYFGFRGGRLDLNGHSLSFHRIQNTDEGAMIV  
 NHNQDKESTVTITGNKDIATGNNNSLDKKEIAYNGWFGEKDTTKTNGRLNLYQPAAEDETRLLLSSGTNLNGNIT  
 QTNGKLFSGRPTPHAYNHLDHWQSKEGI PRGEIWWDNDWINRTFKAENFQIKGGQAVSRNVAVKVGDWHLHSNA  
 QAVFGVAPHQSHTCTRSWTGLTNCKVEKITTDDKVIASLTKTDISGNVDLADAHNLNTGLATLNGNLSANGDTRY  
 TVSHNATQNGNLSLVGNAQATFNOATLNGNTSASGNASFNLSDHAVQNGSLTSGNAKANVSHSALNGNVSADKAV  
 FHFESSRFTGQISGGKDTALHKDSEWTPLSGTELGNNLDNATITLNSAYRHDAAGAQQTGSATDAPRRRSRRSRRS  
 LLSVTPPSTSVERFNTLTNGKLNGQGTFRFMELEFGYRSDKLKLAAESSEGTYTLAVNNTGNEPASLEQLTVVEGKD  
 NKPLSENLNFTLQNEHVDAGAWRYQLIRKDGEFRLHNPKVEQELSDFKLKGKAEEKQAEKDNAQSLDALIAAGRDAVE  
 KTESVAEPARQAGGENVGIMQAEEEEEKRVQADKDTALAKQREAETRPATTAFPRARRARRDLPQLQPOPOPOPQQRDL  
 ISRYANGLSEFSATLNSVFQVQDELDRVFAEDRRNAVWTSGIRDYKHYRSQDFRAYRQTDLRLQIGMQLGNGSGRV  
 GILFSHNRTENTFDDGIGNSARLAHGAVFGQYGIIDRFYIGISAGAGFSSGSLSDGIGGKIRRRVLHYGIQARYRAGF  
 GGFIEPHIGATRYFVQKADYRYENVNIATPGLAFNRYRAGIKADYSFKPAQHISITPYLSYTDAAAGKVRTRVN  
 TAVLAQDFGKTRSAEWGVNAEIKGFTLHAAAAGPQLEAQHSAGIKLGYRW\*

20 To make this construct, the clone pET911 LOmpA (see below) was digested with the *Nhe*I and *Xba*I restriction enzymes and the fragment corresponding to the vector carrying the OmpA leader sequence was purified (pETLOmpA). The ORF1 gene coding for the mature protein was amplified using the oligonucleotides ORF1-For and ORF1-Rev (including the *Nhe*I and *Xba*I restriction sites, respectively), digested with *Nhe*I and *Xba*I and ligated to the purified pETOmpA fragment (see Figure 1). An additional AS dipeptide was introduced by the *Nhe*I site.

25 [0150] All three forms of the protein were expressed. The His-tagged protein could be purified and was confirmed as surface exposed, and possibly secreted (see Figure 3). The protein was used to immunise mice, and the resulting sera gave excellent results in the bactericidal assay.

30 [0151] ORF1LOmpA was purified as total membranes, and was localised in both the inner and outer membranes. Unexpectedly, sera raised against ORF1LOmpA show even better ELISA and anti-bactericidal properties than those raised against the His-tagged protein.

35 [0152] ORF1L was purified as outer membranes, where it is localised.

### 35 Example 7 - protein 911 and its leader peptide

[0153] Protein 911 from *N.meningitidis* (serogroup B, strain MC58) has the following sequence:

40       1 MKKNILEFWV GLFVLIGAAA VAFLAFRVAG GAAFGGSDKT YAVYADFGDI  
 51       51 GGLKVNAPVK SAGVLVGRVG AIGLDPKSYQ ARVRLLDGK YQFSSDVSAQ  
 101     101 ILTSGLLGEQ YIGLQQGGDT ENLAAGDTIS VTSSAMVLEN LIGKFMTSFA  
 151     151 EKNADGGNAE KAAE\*

45 The leader peptide is underlined.

[0154] Three expression strategies have been used for 911:

- 1) 911 with its own leader peptide but without any fusion partner ('911L');
- 2) 911 with the leader peptide from *E.coli*OmpA ('911LOmpA'). To make this construct, the entire sequence encoding the OmpA leader peptide was included in the 5'- primer as a tail (primer 911LOmpA Forward). A *Nhe*I restriction site was inserted between the sequence coding for the OmpA leader peptide and the 911 gene encoding the predicted mature protein (insertion of one amino acid, a serine), to allow the use of this construct to clone different genes downstream the OmpA leader peptide sequence.
- 3) 911 with the leader peptide (MKYLLPTAAAGLLAAQPAMA) from *Erwinia carotovora* PeIB ('911LpeIB').

55 [0155] To make this construct, the 5'-end PCR primer was designed downstream from the leader sequence and included the *Nco*I restriction site in order to have the 911 fused directly to the PeIB leader sequence; the 3'- end primer included the STOP codon. The expression vector used was pET22b+ (Novagen), which carries the coding sequence

for the PeIB leader peptide. The *N*col site introduces an additional methionine after the PeIB sequence.

[0156] All three forms of the protein were expressed. ELISA titres were highest using 911 L, with 919LOmpA also giving good results.

5 **Example 8 - ORF46**

[0157] The complete ORF46 protein from *N.meningitidis* (serogroup B, strain 2996) has the following sequence:

10	1 <u>LGISRKISI</u> LI LSILAVCLPM HAHASDLAND SFIROVLDRQ HFEPDGKYHL
	51   FGRSGELAER SGHIGLGIQ SHQLGNLMIQ QAAIKGNIGY IVRFSDHGHE
	101   VHSFPFDNHAS HSDSDEAGSP VDGFSLYRIH WDGYEHHPAD GYDGPQGGGY
	151   PAPKGARDIY SYDIKGVAQN IRLNLTDRNS TGQRLLADRFH NAGSMLTQGV
	201   GDGFKRATRY SPELDRSGNA AEAFTNGTADI VKNIIGAAGE IVGAGDAVQG
	251   ISEGNSNIAVM HGLGLLSTEN KMARINDLAD MAQLKDYAAA AIRDWAVQNP
15	301   NAAQGIEAVS NIFMAAIPIK GIGAVRGKYG LGGITAHPIK RSQMGAIALP
	351   KGKSAVSDNF ADAAYAKYPY PYHSRNIRSN LEQRYGKENI TSSTVPPSNG
	401   KNVKLADQRH PKTGVPFDGK GFPNFEKHVK YDTKLQEL SGGGIPKAKP
	451   VSDAKPRWEV DRKLNLKLTTR EQVEKVNQEI RNGNKNNSNFS QHAQLEREIN
20	501   KLKSADEINF ADGMGKFTDS MNDAKAFSRLV KSVKENGFTN PVVEYVEING
	551   KAYIVRGNNR VFAAEYLGRI HELKFKKVDF PVPNTSWKNP TDVLNESGNV
	601   KRPRYRSK*

The leader peptide is underlined.

[0158] The sequences of ORF46 from other strains can be found in WO00/66741.

25 [0159] Three expression strategies have been used for ORF46:

- 1) ORF46 with its own leader peptide but without any fusion partner ('ORF46-2L');
- 2) ORF46 without its leader peptide and without any fusion partner ('ORF46-2'), with the leader peptide omitted by designing the 5'-end amplification primer downstream from the predicted leader sequence;

30	1   SDLANDSFIR QVLDRQHFEP DGKYHLFGSR GELAERSGHI GLGKIQSHQL
	51   GNLMIQQAAI KGNIGYIVRF SDHGHEVHSP FDNHASHSDS DEAGSPVDGF
	101   SLYRIHWGDY EHHPADGYDG PQGGGYPAPK GARDIISYDI KGVAQNIRLN
35	151   LTDNRSTGQR LADRPHNAGS MLTQGVGDGF KRATRYSPEL DRSGNAAAEAF
	201   NGTADIVKNI IGAAGEIVGA GDAVQGISEG SNIAVMHGLG LLSTENKMAR
	251   INDLADMAQL KDYAAAIRD WAVQNPNAAO GIEAVSNIFM AAIFIKGIGA
	301   VRGKYGLGGI TAHPIKRSQM GAIALPKGKS AVSDNFADAA YAKYPSPYHS
	351   RNIRSNSLEQR YGKENITSST VPPSNGKNVK LADQRHPKTG VPFDGKGFPN
40	401   FEKHVKYDTK LDIQELSGGG IPKAKPVSDA KPRWEVDRKL NKLTTRREQVE
	451   KNVQEIRNGN KNSNFSQHAQ LEREINKLKS ADEINFADGM GKFTDSMNDK
	501   AFSRLVKSVK ENGFNPVVE YVEINGKAYI VRGNNRVFAA EYLGRIHELK
	551   FKKVDFPVNP TSWKNPTDVL NESGNVKRPR YRSK*

45 3) ORF46 as a truncated protein, consisting of the first 433 amino acids ('ORF46.1L'), constructed by designing PCR primers to amplify a partial sequence corresponding to aa 1-433.  
A STOP codon was included in the 3'-end primer sequences.

50 [0160] ORF46-2L is expressed at a very low level to *E. coli*. Removal of its leader peptide (ORF46-2) does not solve this problem. The truncated ORF46.1L form (first 433 amino acids, which are well conserved between serogroups and species), however, is well-expressed and gives excellent results in ELISA test and in the bactericidal assay.

[0161] ORF46.1 has also been used as the basis of hybrid proteins. It has been fused with 287, 919, and ORF1. The hybrid proteins were generally insoluble, but gave some good ELISA and bactericidal results (against the homologous 2996 strain):

Protein	ELISA	Bactericidal Ab
Orf1-Orf46.1-His	850	256

(continued)

Protein	ELISA	Bactericidal Ab
919-Orf46.1-His	12900	512
919-287-Orf46-His	n.d.	n.d.
Orf46.1-287His	150	8192
Orf46.1-919His	2800	2048
Orf46.1-287-919His	3200	16384

[0162] For comparison, 'triple' hybrids of ORF46.1, 287 (either as a GST fusion, or in  $\Delta$ G287 form) and 919 were constructed and tested against various strains (including the homologous 2996 strain) *versus* a simple mixture of the three antigens. FCA was used as adjuvant:

	2996	BZ232	MC58	NGH38	F6124	BZ133
<b>Mixture</b>	8192	256	512	1024	>2048	>2048
<b>ORF46.1-287-919his</b>	16384	256	4096	8192	8192	8192
<b><math>\Delta</math>G287-919-ORF46.1his</b>	8192	64	4096	8192	8192	16384
<b><math>\Delta</math>G287-ORF46.1-919his</b>	4096	128	256	8192	512	1024

Again, the hybrids show equivalent or superior immunological activity.

[0163] Hybrids of two proteins (strain 2996) were compared to the individual proteins against various heterologous strains:

	1000	MC58	F6124 (MenA)
<b>ORF46.1-His</b>	<4	4096	<4
<b>ORF1-His</b>	8	256	128
<b>ORF1—ORF46.1-His</b>	1024	512	1024

[0164] Again, the hybrid shows equivalent or superior immunological activity.

#### Example 9 - protein 961

[0165] The complete 961 protein from *N.meningitidis* (serogroup B, strain MC58) has the following sequence:

```

1  MSMKHFPAKV LTTAILATFC SGALAATSDD DVKKAATVAI VAAYNNNGQEI
51 NGFKAGETIY DIGEDGTITQ KDATAADVEA DDFKGIGLKK VVTNLTKTVN
101 ENKQNVDAKV KAAESEIEKL TTKLADTDAA LADTDAALDE TTNALNKLGE
151 NITTFAEETK TNIVKIDEKL EAVADTVDKH AEAFNDIADS LDETNTKADE
201 AVKTANEAKQ TAEETKQNVD AKVKAATCAA GKAEEAAGTA NTAADKAEAV
251 AAKVTDIKAD IATNKADIAK NSARIDSLSDK NVANLRKETR QGLAEQAALS
301 GLFQPYNVGR FNVTAAVGGY KSESAVAIGT GFRFTENFAA KAGVAVGTSS
351 GSSAAYHVGV NYEW*

```

[0166] The leader peptide is underlined.

[0167] Three approaches to 961 expression were used:

- 55 1) 961 using a GST fusion, following WO99/57280 ('GST961');
- 2) 961 with its own leader peptide but without any fusion partner ('961L'); and
- 3) 961 without its leader peptide and without any fusion partner ('961<sup>untagged</sup>'), with the leader peptide omitted by

designing the 5'-end PCR primer downstream from the predicted leader sequence.

[0168] All three forms of the protein were expressed. The GST-fusion protein could be purified and antibodies against it confirmed that 961 is surface exposed (Figure 4). The protein was used to immunise mice, and the resulting sera gave excellent results in the bactericidal assay. 961 L could also be purified and gave very high ELISA titres.

[0169] Protein 961 appears to be phase variable. Furthermore, it is not found in all strains of *N.meningitidis*.

**Example 10 - protein 287**

[0170] Protein 287 from *N.meningitidis* (serogroup B, strain 2996) has the following sequence:

```

1  MFERSVIAMA CIFALSACGG GGGGSPDVKS ADTLSKPAAP VVAEKETEVK
51  EDAPQAGSQG QGAPSTQGSQ DMAAVSAENT GNNGAATTDK PKNEDEGPQN
101 DMPQNSAESA NQTGNNQPAD SSDSAPASNP APANGGSNFG RVDLANGVLI
151 DGPSQNITLT HCKGDSCNGD NLLDEEAPSK SEFENLNESE RIEKYKKDGK

20  SDKFTNLVAT AVQANGTNKY VIIYKDKSAS SSSARFRRSA RSRRSLPAEM
251 PLIPVNQADT LIVDGEAVSL TGHSGNIFAP EGNYRYLTYG AEKLPGGSYA
301 LRVQGEPAKG EMLAGTAVYN GEVLHFHTEN GRPYPTRGRF AAKVDFGSKS
351 VDGIIDSGDD LHMGTQKFKA AIDGNGFKGT WTENGGGDVS GRFYGPAGEEE
401 VAGKYSYRPT DAEKGFFGVF AGKKEQD*

```

[0171] The leader peptide is shown underlined.

[0172] The sequences of 287 from other strains can be found in Figures 5 and 15 of WO00/66741.

[0173] Example 9 of WO99/57280 discloses the expression of 287 as a GST-fusion in *E.coli*.

[0174] A number of further approaches to expressing 287 in *E.coli* have been used, including:

- 30 1) 287 as a His-tagged fusion ('287-His');
- 2) 287 with its own leader peptide but without any fusion partner ('287L');
- 3) 287 with the ORF4 leader peptide and without any fusion partner ('287LOrf4');
- 4) 287 without its leader peptide and without any fusion partner ('287<sup>untagged</sup>');

```

1  CGGGGGGSPD VKSADTLSKP AAPVVAEKET EVKEDAPQAG SQGQGAPSTQ
51  GSQDMAAVSA ENTGNGGAAT TDKPKNEDEG PQNDMPQNSA ESANQTGNNQ
101 PADSSDAPA SNPAPANGGS NFGRVDSLNG VLIDGPSQNI TLTHCKGDSC
151 NGDNLLDEEA PSKSEFENLN ESERIEKYKK DGKSDKFTNL VATAVQANGT
201 NKYIIYKDK SASSSSARFR RSARSRRSLP AEMPLIPVNO ADTLIVDGEA
251 VSLTGHSgni FAPEGNYRL TYGAEKLPGG SYALRVQGEP AKGEMLAGTA
301 VYNGEVILFH TENGRPYPTR GRFAAKVDFG SKSVDGIIDS GDDLHMGQTQK
351 FKAIAIDGNGF KGTWTENGGG DVSGRFYGPAA GEEVAGKSY RPTDAEKGGF
401 GVFAGKKEQD *

```

[0175] All these proteins could be expressed and purified.

[0176] '287L' and '287LOrf4' were confirmed as lipoproteins.

[0177] As shown in Figure 2, '287LOrf4' was constructed by digesting 919LOrf4 with *Nhe*I and *Xho*I. The entire ORF4 leader peptide was restored by the addition of a DNA sequence coding for the missing amino acids, as a tail, in the 5'-end primer (287LOrf4 for), fused to 287 coding sequence. The 287 gene coding for the mature protein was amplified using the oligonucleotides 287LOrf4 For and Rev (including the *Nhe*I and *Xho*I sites, respectively), digested with *Nhe*I and *Xho*I and ligated to the purified pETOorf4 fragment.

**Example 11 - further non-fusion proteins with/without native leader peptides**

[0178] A similar approach was adopted for *E.coli* expression of further proteins from WO99/24578, WO99/36544 and WO99/57280.

[0179] The following were expressed without a fusion partner: 008, 105, 117-1, 121-1, 122-1, 128-1, 148, 216, 243, 308, 593, 652, 726, 982, and Orf143-1. Protein 117-1 was confirmed as surface-exposed by FACS and gave high ELISA titres.

[0180] The following were expressed with the native leader peptide but without a fusion partner: 111, 149, 206, 225-1, 235, 247-1, 274, 283, 286, 292, 401, 406, 502-1, 503, 519-1, 525-1, 552, 556, 557, 570, 576-1, 580, 583, 664, 759, 907, 913, 920-1, 926, 936-1, 953, 961, 983, 989, Orf4, Orf7-1, Orf9-1, Orf23, Orf25, Orf37, Orf38, Orf40, Orf40.1, Orf40.2, Orf72-1, Orf76-1, Orf85-2, Orf91, Orf97-1, Orf119, Orf143.1. These proteins are given the suffix 'L'.

[0181] His-tagged protein 760 was expressed with and without its leader peptide. The deletion of the signal peptide greatly increased expression levels. The protein could be purified most easily using 2M urea for solubilisation.

[0182] His-tagged protein 264 was well-expressed using its own signal peptide, and the 30kDa protein gave positive Western blot results.

[0183] All proteins were successfully expressed.

[0184] The localisation of 593, 121-1, 128-1, 593, 726, and 982 in the cytoplasm was confirmed.

[0185] The localisation of 920-1L, 953L, ORF9-1L, ORF85-2L, ORF97-1L, 570L, 580L and 664L in the periplasm was confirmed.

[0186] The localisation of ORF40L in the outer membrane, and 008 and 519-1L in the inner membrane was confirmed. ORF25L, ORF4L, 406L, 576-1L were all confirmed as being localised in the membrane.

[0187] Protein 206 was found not to be a lipoprotein.

[0188] ORF25 and ORF40 expressed with their native leader peptides but without fusion partners, and protein 593 expressed without its native leader peptide and without a fusion partner, raised good anti-bactericidal sera. Surprisingly, the forms of ORF25 and ORF40 expressed without fusion partners and using their own leader peptides (i.e. 'ORF25L' and 'ORF40L') give better results in the bactericidal assay than the fusion proteins.

[0189] Proteins 920L and 953L were subjected to N-terminal sequencing, giving HRVWVETAH and ATYKVDEY-HANARFAF, respectively. This sequencing confirms that the predicted leader peptides were cleaved and, when combined with the periplasmic location, confirms that the proteins are correctly processed and localised by *E.coli* when expressed from their native leader peptides.

[0190] The N-terminal sequence of protein 519.1L localised in the inner membrane was MEFFIILLA, indicating that the leader sequence is not cleaved. It may therefore function as both an uncleaved leader sequence and a transmembrane anchor in a manner similar to the leader peptide of PBP1 from *N.gonorrhoeae* [Ropp & Nicholas (1997) J. Bact. 179: 2783-2787]. Indeed the N-terminal region exhibits strong hydrophobic character and is predicted by the Tmpred. program to be transmembrane.

#### **Example 12 - lipoproteins**

[0191] The incorporation of palmitate in recombinant lipoproteins was demonstrated by the method of Kraft et. al. [J. Bact. (1998) 180:3441-3447.]. Single colonies harbouring the plasmid of interest were grown overnight at 37°C in 20 ml of LB/Amp (100µg/ml) liquid culture. The culture was diluted to an OD<sub>550</sub> of 0.1 in 5.0 ml of fresh medium LB/Amp medium containing 5 µC/ml [<sup>3</sup>H] palmitate (Amersham). When the OD<sub>550</sub> of the culture reached 0.4-0.8, recombinant lipoprotein was induced for 1 hour with IPTG (final concentration 1.0 mM). Bacteria were harvested by centrifugation in a bench top centrifuge at 2700g for 15 min and washed twice with 1.0 ml cold PBS. Cells were resuspended in 120µl of 20 mM Tris-HCl (pH 8.0), 1 mM EDTA, 1.0% w/v SDS and lysed by boiling for 10 min. After centrifugation at 13000g for 10 min the supernatant was collected and proteins precipitated by the addition of 1.2 ml cold acetone and left for 1 hour at -20 °C. Protein was pelleted by centrifugation at 13000g for 10 min and resuspended in 20-50µl (calculated to standardise loading with respect to the final O.D of the culture) of 1.0% w/v SDS. An aliquot of 15 µl was boiled with 5µl of SDS-PAGE sample buffer and analysed by SDS-PAGE. After electrophoresis gels were fixed for 1 hour in 10% v/v acetic acid and soaked for 30 minutes in Amplify solution (Amersham). The gel was vacuum-dried under heat and exposed to Hyperfilm (Kodak) overnight -80 °C.

[0192] Incorporation of the [<sup>3</sup>H] palmitate label, confirming lipidation, was found for the following proteins: Orf4L, Orf25L, 287L, 287LOrf4, 406.L, 576L, 926L, 919L and 919LOrf4.

#### **Example 13 - domains in 287**

[0193] Based on homology of different regions of 287 to proteins that belong to different functional classes, it was split into three 'domains', as shown in Figure 5. The second domain shows homology to IgA proteases, and the third domain shows homology to transferrin-binding proteins.

[0194] Each of the three 'domains' shows a different degree of sequence conservation between *N.meningitidis* strains - domain C is 98% identical, domain A is 83% identical, whilst domain B is only 71% identical. Note that protein 287 in strain MC58 is 61 amino acids longer than that of strain 2996. An alignment of the two sequences is shown in Figure

7, and alignments for various strains are disclosed in WO00/66741 (see Figures 5 and 15 therein).

[0195] The three domains were expressed individually as C-terminal His-tagged proteins. This was done for the MC58 and 2996 strains, using the following constructs:

287a-MC58 (aa 1-202), 287b-MC58 (aa 203-288), 287c-MC58 (aa 311-488).

5 287a-2996 (aa 1-139), 287b-2996 (aa 140-225), 287c-2996 (aa 250-427).

[0196] To make these constructs, the stop codon sequence was omitted in the 3'-end primer sequence. The 5' primers included the *Nhel* restriction site, and the 3' primers included a *Xhol* as a tail, in order to direct the cloning of each amplified fragment into the expression vector pET21 b+ using *Ndel-Xhol*, *Nhel-Xhol* or *Ndel-HindIII* restriction sites.

[0197] All six constructs could be expressed, but 287b-MC8 required denaturation and refolding for solubilisation.

10 [0198] Deletion of domain A is described below ('Δ4 287-His').

[0199] Immunological data (serum bactericidal assay) were also obtained using the various domains from strain 2996, against the homologous and heterologous MenB strains, as well as MenA (F6124 strain) and MenC (BZ133 strain):

	<b>2996</b>	<b>BZ232</b>	<b>MC58</b>	<b>NGH38</b>	<b>394/98</b>	<b>MenA</b>	<b>MenC</b>	
	<b>287-His</b>	32000	16	4096	4096	512	8000	16000
	<b>287(B)-His</b>	256	-	-	-	-	16	-
	<b>287(C)-His</b>	256	-	32	512	32	2048	>2048
15	<b>287(B-C)-His</b>	64000	128	4096	64000	1024	64000	32000
20								

[0200] Using the domains of strain MC58, the following results were obtained:

	<b>MC58</b>	<b>2996</b>	<b>BZ232</b>	<b>NGH38</b>	<b>394/98</b>	<b>MenA</b>	<b>MenC</b>	
	<b>287-His</b>	4096	32000	16	4096	512	8000	16000
	<b>287(B)-His</b>	128	128	-	-	-	-	128
	<b>287(C)-His</b>	-	16	-	1024	-	512	-
25	<b>287(B-C)-His</b>	16000	64000	128	64000	512	64000	>8000
30								

#### **Example 14 — deletions in 287**

[0201] As well as expressing individual domains, 287 was also expressed (as a C-terminal His-tagged protein) by making progressive deletions within the first domain. These

[0202] Four deletion mutants of protein 287 from strain 2996 were used (Figure 6):

- 40 1) '287-His', consisting of amino acids 18-427 (*i.e.* leader peptide deleted);
- 2) 'Δ1 287-His', consisting of amino acids 26-427;
- 3) 'Δ2 287-His', consisting of amino acids 70-427;
- 4) 'Δ3 287-His', consisting of amino acids 107-427; and
- 5) 'Δ4 287-His', consisting of amino acids 140-427 (=287-bc).

45 [0203] The 'Δ4' protein was also made for strain MC58 ('Δ4 287MC58-His'; aa 203-488).

[0204] The constructs were made in the same way as 287a/b/c, as described above.

[0205] All six constructs could be expressed and protein could be purified. Expression of 287-His was, however, quite poor.

[0206] Expression was also high when the C-terminal His-tags were omitted.

50 [0207] Immunological data (serum bactericidal assay) were also obtained using the deletion mutants, against the homologous (2996) and heterologous MenB strains, as well as MenA (F6124 strain) and MenC (BZ133 strain):

	<b>2996</b>	<b>BZ232</b>	<b>MC58</b>	<b>NGH38</b>	<b>394/98</b>	<b>MenA</b>	<b>MenC</b>	
	<b>287-his</b>	32000	16	4096	4096	512	8000	16000
	<b>Δ1 287-His</b>	16000	128	4096	4096	1024	8000	16000
55	<b>Δ2 287-His</b>	16000	128	4096	>2048	512	16000	>8000

(continued)

	2996	BZ232	MC58	NGH38	394/98	MenA	MenC
5	Δ3 287-His	16000	128	4096	>2048	512	16000
	Δ4 287-His	64000	128	4096	64000	1024	64000
							32000

[0208] The same high activity for the Δ4 deletion was seen using the sequence from strain MC58.

[0209] As well as showing superior expression characteristics, therefore, the mutants are immunologically equivalent or superior.

#### Example 15 - poly-glycine deletions

[0210] The 'Δ1 287-His' construct of the previous example differs from 287-His and from '287<sup>untagged</sup>' only by a short N-terminal deletion (GGGGGGS). Using an expression vector which replaces the deleted serine with a codon present in the *Nhe* cloning site, however, this amounts to a deletion only of (Gly)<sub>6</sub>. Thus, the deletion of this (Gly)<sub>6</sub> sequence has been shown to have a dramatic effect on protein expression.

[0211] The protein lacking the N-terminal amino acids up to GGGGGG is called 'ΔG 287'. In strain MC58, its sequence (leader peptide underlined) is:

ΔG287

1	MFKRSVIAMA	CIFALSACGG	GGGGSPDVKS	ADTLSKPAAP	VVSEKETEAK
51	EDAPQAGSQG	QGAPSAQGSQ	DMAAVSEENT	GNGGAVTADN	PKNEDDEVAQN
101	DMPQNAAGTD	SSTPNHTPDP	NMLAGNMENQ	ATDAGESSQP	ANQPDMANAA
151	DGMQGDDPSA	GGQNAGNTAA	QGANQAGNNQ	AAGSSDFIPA	SNPAPANGGS
201	NFGRVDSLNG	VLIDGPSQNI	TLTHCKGDSC	SGNNFLDEEV	QLKSEEFKLS
251	DADKISNYKK	DGKNDKFVG	VADSVQMKG	NQYIIFYKPK	PTSFARFRRS
301	ARSRRSLPAE	MPLIPVNQAD	TLIVDGEAVS	LTHSGNIFA	PEGNYRYLTY
351	GAEKLPGGSY	ALRVQGEPAK	GEMLAGAAVY	NGEVLHFHTE	NGRPYPTRGR
401	FAAKVDFGSK	SVDGIIDSQD	DLHMGTQKF	AAIDGNGFKG	TWTENGSGDV
451	SGKFYGPAGE	EVAGKYSYRP	TDAEKGGFGV	FAGKKEQD*	

[0212] ΔG287, with or without His-tag ('ΔG287-His' and 'ΔG287K', respectively), are expressed at very good levels in comparison with the '287-His' or '287 untagged'.

[0213] On the basis of gene variability data, variants of ΔG287-His were expressed in *E.coli* from a number of MenB strains, in particular from strains 2996, MC58, 1000, and BZ232. The results were also good.

[0214] It was hypothesised that poly-Gly deletion might be a general strategy to improve expression. Other MenB lipoproteins containing similar (Gly)<sub>n</sub> motifs (near the N-terminus, downstream of a cysteine) were therefore identified, namely Tbp2 (NMB0460), 741 (NMB 1870) and 983 (NMB1969):

ΔGTbp2

1	MNNPLVNQAA	MVLPVFLLSA	CLGGGGSF	DSVDT	TEAPRP	APKYQDV	FSE	
51	KPQAQKDQGG	YGFAMRLKRR	NWYPQAK	EDEV	VKL	DESDWEA	TGLP	DPEPKEL
101	PKRQKSVIEK	VET	DSNDNNI	YSSPYLKPSNH	QNGNT	GNGIN	QPKNQAKDYE	
151	NFKYVYSGWF	YKHAKREFNL	KVEPKSAKNG	DDGYI	FYHGK	EPSRQLPASG		
201	KITYKGWHF	ATDTKKGQKF	REIIQPSKSQ	GDRYSGFSGD	DGE	EYSNK	DNK	
251	STLTDGQEGY	GFTSNLEVDF	HNKKLTGKLI	RNNANTDNNQ	ATT	TQYYSLE		
301	AQVTGNRFNG	KATATDKPQQ	NSETKEHPFV	SDSSSLGGF	FGPQGEELGF			
351	RFLSDDQKVA	VVGSAKTKDK	PANGNTAAAS	GGTDAASNG	AAGTSSEN	GK		
401	LTTVLD	DAVEL	KLGDKEVQKL	DNFSNAAQLV	VDGIMIPLL	EASESGNNQ	A	

451 NOGTNGGTAF TRKFDHTPES DKKDAQAGTQ TNGAQTASNT AGDTNGKTKT  
 501 YEVEVCCSNL NYLKYGMLTR KNSKSAMQAG ESSSQADAKT EQVEQSMFLQ  
 551 GERTDEKEIP SEQNIVYRGS WYGYIANDKS TSWSGNASNA TSGNRAEFTV  
 5 NFADKKITGT LTADNRQEAT FTIDGNIKDN GFEGTAKTAE SGFDLDQSNT  
 651 TRTPKAYITD AKVQGGFYGP KAEELGGWFA YPGDKQTKNA TNASGNSSAT  
 701 VVFGAKRQQP VR\*

10 741 / ΔG741  
 1 VNRTAFCCLS LTTALIILTAC SSGGGGVAAAD IAGGLADALT APLDHDKGL  
 51 QSLTLDQSVR KNEKLKLAQ GAETKTYGNGD SLNTGKLKD KVSRDFD FIRQ  
 101 IEVDGQLITL ESGEFQVYKQ SHSALTAFQT EQIQDSEHSG KMVAKRQFRI  
 151 GDIAGEHTSF DKLPEGGRAT YRGTAFGSDD AGGKLTYTID FAAKQGNGKI  
 201 EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA  
 251 QEVARSAEVK TVNGIRHIGL AAKQ\*

20 983 / ΔG983  
 1 MRTTPTFPTK TFKPTAMALA VATTLSACLG GGGGGTSAPD FNAGGTGIGS  
 51 NSRATTAKSA AVSYAGIKNE MCKDRSMLCA GRDDVAVTDR DAKINAPPN  
 101 LHTGDFPNPN DAYKNLINLK PAIEAGYTGR GVEVGIVDTG ESVGSISFPE  
 151 LYGRKEHGYN ENYKNYTAYM RKEAPEDGGG KDIEASFDE AVIETEAKPT  
 201 DIRHVKEIGH IDLVSHI1GG RSVDGRPAGG IAPDATLHIM NTNDETKNEM  
 251 MVAAIRNAWV KLGERGVRIV NNSFGTTSRA GTADLFQIAN SEEQYROALL  
 301 DYSGGDKTDE GIRLMQOSDY GNLSYHIRNK NMLFISTGN DAQAQPN TYA  
 351 LLPFYEKDAQ KGIITVAGVD RSGEKFKREM YGEPGTEPLE YGSNHCGITA  
 401 MWCLSAPYEA SVRFTRTNPI QIAGTSFSAP IVTGTAA ALL QKYPWMSNDN  
 451 LRTTLLTTAQ DIGAVGVDSK FGWGLLDAGK AMNGPASFPF GDFTADTKGT  
 501 SDIAYSFRND ISGTGGLIKK GGSQQLQHGN NTYTGKTIIE GGSLVLYGNN  
 551 KSDMRVETKG ALIYNGAASG GSLSNDGIVY LADTDQSGAN ETVHIKGSLQ  
 601 LDGKGTLTYTR LGKLLKVDGT AIIGGGKLYMS ARGKGAGYLN STGRRVPFLS  
 651 AAKIGQDYSF FTNIETDGGL LASLDSVEKT AGSEGDTLSY YVRGNAART  
 701 ASAAAHSAPA GLKHAVEQGG SNLENLNMVEL DASESSATPE TVETAAADRT  
 751 DMMPGIRPYGA TFRAAAAVQH ANAADGVRIF NSLAATVYAD STAHAADMQG  
 801 RRLKAVSDGL DHNGTGLRVI AQTQQDGGTW EQGGVEGKMR GSTQTVGIAA  
 851 KTGEINTTAA TLGMGRSTWS ENSANAKTDS ISLFAGIRHD AGDIGYLKGL  
 901 FSYGRYKNIS SRSTGADEHA EGSVNGTLMQ LGALGGVNVP FAATGDLTVE  
 951 GGLRYDLLKQ DAFAEKGSAL GWGNSNLTEG TLVGLAGLKL SQPLSDKAVL  
 1001 FATAGVERDL NGRDVTVTGG FTGATAATGK TGARNMPHTR LVAGLGADVE  
 1051 FGNGWNGLAR YSYAGSKQYG NHSGRVGVGY RF\*

40 [0215] Tbp2 and 741 genes were from strain MC58; 983 and 287 genes were from strain 2996. These were cloned in pET vector and expressed in *E.coli* without the sequence coding for their leader peptides or as "ΔG forms", both fused to a C-terminal His-tag. In each case, the same effect was seen - expression was good in the clones carrying the deletion of the poly-glycine stretch, and poor or absent if the glycines were present in the expressed protein:

	ORF	Express.	Purification	Bact. Activity
45	287-His(2996)	+/-	+	+
50	'287untagged'(2996)	+/-	nd	nd
	ΔG287-His(2996)	+	+	+
	ΔG287K(2996)	+	+	+
	ΔG287-His(MC58)	+	+	+
	ΔG287-His(1000)	+	+	+
	ΔG287-His(BZ232)	+	+	+
55	Tbp2-His(MC58)	+/-	nd	nd
	ΔGTbp2-His(MC58)	+	+	
	741-His(MC58)	+/-	nd	nd

(continued)

ORF	Express.	Purification	Bact. Activity
ΔG741-His(MC58)	+	+	
983-His (2996)			
ΔG983-His (2996)	+	+	

[0216] SDS-PAGE of the proteins is shown in Figure 13.

*ΔG287 and hybrids*

[0217] ΔG287 proteins were made and purified for strains MC58, 1000 and BZ232. Each of these gave high ELISA titres and also serum bactericidal titres of >8192. ΔG287K, expressed from pET-24b, gave excellent titres in ELISA and the serum bactericidal assay. ΔG287-ORF46.1K may also be expressed in pET-24b.

[0218] ΔG287 was also fused directly in-frame upstream of 919, 953, 961 (sequences shown below) and ORF46.1:

ΔG287-919

1	ATGGCTAGCC CCGATGTTAA ATCGGCGGAC ACGCTGTCAA AACCGGCCGC
51	TCCCTGTTGTT GCTGAAAAAG AGACAGAGGT AAAAGAAGAT GCGCCACAGG
101	CAGGTTCTCA AGGACAGGGC GCGCCATCCA CACAAGGCAG CCAAGATATG
151	GCGGCAGTTT CGGCAGAAAA TACAGGCAAT GGCGGTGCGG CAACAACGGA
201	CAAACCCAAA AATGAAGACG AGGGACCGCA AAATGATATG CCGCAAAATT
251	CCGCCGAATC CGCAAATCAA ACAGGGAACA ACCAACCCGC CGATTCTTCA
301	GATTCCGCC CCGCGTCAAA CCCTGCACCT GCGAATGGCG GTAGCAATT
351	TGGAAGGGTT GATTGGCTA ATGGCGTTT GATTGATGGG CCGTCGCAAA
401	ATATAACGTT GACCCACTGT AAAGGCGATT CTTGTAATGG TGATAATT
451	TTGGATGAAG AAGCACCGTC AAAATCAGAA TTTGAAAATT TAAATGAGTC
501	TGAACGAATT GAGAAATATA AGAAAAGATGG GAAAAGCGAT AAATTTACTA
551	ATTTGGTTGC GACAGCAGTT CAAGCTAATG GAACTAACAA ATATGTCATC
601	ATTTATAAAAG ACAAGTCCGC TTCATCTTCA TCTGCGCGAT TCAGGCCTTC
651	TGCACGGTCG AGGAGGTCGC TTCCCTGCCGA GATGCCGCTA ATCCCCGTCA
701	ATCAGGCCGA TACGCTGATT GTCGATGGGG AAGCGGTGAG CCTGACGGGG
751	CATTCCGGCA ATATCTTCGC GCCCGAAGGG AATTACCGGT ATCTGACTTA
801	CGGGGCGGAA AAATGCCCG GCGGATCGTA TGCCCTCCGT GTGCAAGGCG
851	AACCGGCAAA AGGCAGAAATG CTTGCTGCCA CGGCCGTGTA CAAACGGCGAA
901	GTGCTGCATT TTCATACGGA AAACGGCCGT CCGTACCCGA CTAGAGGCAG
951	GTTTGGCGCA AAAGTCGATT TCGGCAGCAA ATCTGTGGAC GGCATTATCG
1001	ACAGCGGCCGA TGATTGCGAT ATGGGTACGC AAAAATTCAA AGCCGCCATC

20

15

25

30

35

40

45

50

55

	1051	GATGGAAACG GCTTTAAGGG GACTTGGACG GAAAATGGCG GCGGGGATGT
5	1101	TTCCCGGAAGG TTTTACGGCC CGGCCGGCG GGAAGTGGCG GGAAAATACA
	1151	GCTATCGCCC GACAGATGCG GAAAAGGGCG GATTCCGCCT GTTTGCCGGC
	1201	AAAAAAAGAGC AGGATGGATC CGGAGGAGGA GGATGCCAAA GCAAGAGCAT
	1251	CCAAACCTT CCGCAACCCG ACACATCCGT CATCAACGGC CCGGACCGGC
	1301	CGGTGGCAT CCCCAGCCCC CGCGAACCGA CGGTGGCGG CGGCAGGGGCC
	1351	GTCTATAACCG TTGTAACCGA CCTGTCCCTG CCCCACTGGG CGGCGCAGGA
	1401	TTTCGCCAAA AGCCTGCAAT CCTTCCGCCT CGGCTGCGCC AATTGAAAAA
10	1451	ACCGCCAAGG CTGGCAGGAT TGTCGCGGCC AAGCCTTCA AACCCCCGTC
	1501	CATTCTTTC AGGCAAAACCA GTTTTTGAA CGCTATTCA CGCCGTGGCA
	1551	GGTTGCAGGC AACGGAAGCC TTGCGGTAC GGTTACCGGC TATTACGAGC
	1601	CGGTGCTGAA GGGCGACGAC AGGCGGGACGG CACAAGCCCG CTTCGGATT
	1651	TACGGTATTC CCGACGATTT TATCTCCGTC CCCCTGCCTG CCGGTTTGCG
	1701	GAGCGGAAAA GCCCTTGTCC GCATCAGGCA GACGGGAAAA AACAGCGGC
15	1751	CAATCGACAA TACCGGCGGC ACACATACCG CCGACCTCTC CCGATTCCCC
	1801	ATCACCGCGC GCACAAACGGC AATCAAAGGC AGGTTTGAAAG GAAGCCGCTT
	1851	CCTCCCCCTAC CACACCGCGA ACCAAATCAA CGGCGCGCG CTTGACGGCA
	1901	AAGCCCCGAT ACTCGGTTAC CGCGAAGACC CGTCTGAAC TTTTTTTATG
	1951	CACATCCAAG GCTCGGGCCG TCTGAAAACC CGTCTCGGC AATACATCCG
20	2001	CATCGGCTAT GCCGACAAAA ACGAACATCC CTACGTTCC ATCGGACGCT
	2051	ATATGGCGGA CAAAGGCTAC CTCAAGCTCG GGCAGACCTC GATGCAGGGC
	2101	ATCAAAGCCT ATATCGGGCA AAATCCGCAA CGCCTCGCCG AAGTTTTGGG
	2151	TCAAAACCCC AGCTATATCT TTTTCCGCGA GCTTGGCGGA AGCAGCAATG
	2201	ACGGTCCCGT CGGCGCACTG GGCACCCCGT TGATGGGGGA ATATGCCGGC
	2251	GCAGTCGACC GGCACTACAT TACCTTGGGC GCGCCCTTAT TTGTCGCCAC
	2301	CGCCCATCCG GTTACCCGCA AAGCCCTCAA CGCCTGATT ATGGCGCAGG
25	2351	ATACCGGCAG CGCGATTAAA GGCGCGGTGC GCGTGGATTA TTTTTGGGGA
	2401	TACGGCGACG AAGCCGGCGA ACTTGCCGGC AAACAGAAAA CCACGGGTTA
	2451	CGTCTGGCAG CTCCTACCCA ACGGTATGAA GCGCGAATAC CGCCCGTAAC
	2501	TCGAG
30	1	MASPDVKSAD TLSKPAAPVV AEKETEVKED APQAGSQGQG APSTQGSQDM
	51	AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNNQPADSS
	101	DSAPASNAP ANGGSNFGRV DLANGVLIDG PSQNITLTHC KGDSCNGDNL
35	151	LDEEAPSKSE FENLINESERI EKYKKDGKSD KFTNLVATAV QANGTNKYVI
	201	IYKDKSASSS SARFRRSARS RRSRPAEMPL IPVNQADTLI VDGEAVSLTG
	251	HSGNIFAPEG NYRYLTYGAE KLPGGSYALR VQGEPAKgem LAGTAVYNge
	301	VLHFHTENGR PYPTGRFAA KVDFGSKSVD GIIDSGDDLH MGTQKFKAII
	351	DGNGFKGTWT ENGGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGFFGVFAG
	401	KKEQDGSGGG GCQSFSIQTF PQPDTSVING PDRPVGIPDP AGTTVGGGGAA
	451	VYTVPVPHSL PHWAAQDFAK SLQSFRRLGCA NLKNRQGWQD VCAQAFQTPV
40	501	HSFQAKQFFE RYFTPWQVAG NGSLAGTVTC YYEPVLIKGD RRTAQARFPI
	551	YGIPDDFISV PLPAGLRSRK ALVRIRQTK NSGTIDNTGG THADLSRFP
	601	ITARTTAIKG RFEGRFLPY HTRNQINGGA LDGKAPILY AEDPVELFFM
	651	HIQGSGRSLKT PSGKYIRIGY ADKNEHPYVS IGRYMADKGY LKLGQTSMQG
	701	IKAYMRQNPQ RLAEVLGQNP SYIFFRELAG SSNDGPVGAL GTPLMGEYAG
	751	AVDRHYITLG APLFVATAHP VTRKALNRLI MAQDTGSAIK GAVRVDYFWG
45	801	YGDEAGELAG KQKTTGYVWQ LLPNGMKPEY RP*

AG287-953

5            1 ATGGCTAGCC CCGATGTTAA ATCGGGCGAC ACGCTGTCAA AACCGGCCGC  
       51 TCCTGTTGTT GCTGAAAAAG AGACAGAGGT AAAAGAAGAT GCGCCACAGG  
       101 CAGGTTCTCA AGGACAGGGC CGGCCATCCA CACAAGGCAG CCAAGATATG  
       151 GCGGCAGTTT CGGCAGAAAA TACAGGCAAT GCGGGTGCAG CAACAACGGA  
       201 CAAACCCAAA AATGAAGACG AGGGACCGCA AAATGATATG CCGCAAAATT  
       251 CCGCGAATC CGCAAATCAA ACAGGGAAACA ACCAACCCGC CGATTCTTC  
       301 GATTCCGCC CGCGTCAAA CCCTGCACCT GCGAATGGCG GTAGCAATT  
       351 TGGAAGGGTT GATTGGCTA ATGGCTTTT GATTGATGGG CCGTCGAAA  
 10            401 ATATAACGTT GACCACTGT AAAGGCAGATT CTTGTAATGG TGATAATTAA  
       451 TTGGATGAAAG AAGCACCGTC AAAATCAGAA TTTGAAAATT TAAATGAGTC  
       501 TGAACGAATT GAGAAATATA AGAAAGATGG GAAAAGCGAT AAATTTACTA  
       551 ATTTGGTGTG GACAGCAGTT CAAGCTAATG GAACTAACAA ATATGTCATC  
       601 ATTTATAAAG ACAAGTCCC TTCATCTTC TCTGCCGAT TCAGGCCGTT  
       651 TGCACGGTGC AGGAGGTCGC TCCCTGCCA GATGCCGCTA ATCCCCGTCA  
 15            701 ATCAGGGGA TACGCTGATT GTCGATGGGG AAGCGGTCAG CCTGACGGGG  
       751 CATTCCGGCA ATATCTCGC GCCCGAAGGG AATTACCGGT ATCTGACTTA

20            801 CGGGGCGGAA AAATTGCCCG GCGGATCGTA TGCCCTCCGT GTGCAAGGCG  
       851 AACCGGCAAA AGGCGAAATG CTTGCTGGCA CGGCCGTGTA CAACGGCGAA  
       901 GTGCTGCATT TTCATACCGA AAACGGCCGT CGTACCCGA CTAGAGGCAG  
       951 GTTGGCGCA AAAGTCGATT TCGGCAGCAA ATCTGTGGAC GGCATTATCG  
 25            1001 ACAGCGGCGA TGATTTGCAT ATGGGTACCG AAAAATCAA AGCCGCCATC  
       1051 GATGGAACG GCTTTAAGGG GACTTGGACG GAAAATGGCG GCGGGGATGTT  
       1101 TTCCGGAAGG TTTTACGGCC CGGCCGGCGA GGAAGTGGCG GAAAAATACAA  
       1151 GCTATGCCCG GACAGATGGCG GAAAAGGGCG GATTCCGGGT GTTTGCCGGC  
       1201 AAAAAAGAGC AGGATGGATC CGGAGGAGGA GGAGCCACCT ACAAAAGTGG  
       1251 CGAATATCAC GCCAACGCCCG GTTTCGCCAT CGACCATTTC AACACCAGCA  
 30            1301 CCAACGTCGG CGGTTTTAC GGTCTGACCG GTTCCGTCGA GTTCGACCAA  
       1351 GCAAAACGCG ACGGTAAAAT CGACATCACC ATCCCCGTTG CCAACCTGCA  
       1401 AAGCGGTTCG CAACACTTTA CCGACCACCT GAAATCAGCC GACATCTTC  
       1451 ATGCCGCCA ATATCCGGAC ATCCGCTTTG TTTCCACCAA ATTCAACTTC  
       1501 AACGGAAAAA AACTGGTTTC CTTTGACGCG AACCTGACCA TGCACGGCAA  
 35            1551 AACCGCCCCC GTCAAACCTCA AAGCCGAAAA ATTCAACTG TACCAAAGGC  
       1601 CGATGGCGAA AACCGAAGTT TGCGGGCGCG ACTTCAGCAC CACCATCGAC  
       1651 CGCACCAAAAT GGGCGTGGCA CTACCTCGT AACGTTGGTA TGACCAAAAG  
       1701 CGTCCGCATC GACATCCAAA TCGAGGCAGC CAAACAATAA CTCGAG

40            1 MASPDVKSAD TLSKPAAPVV AEKETEVKED APQAGSQGQG APSTQGSQDM  
       51 AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNNQPADSS  
       101 DSAPASNPAP ANGGSNFGRV DLANGVLIDG PSQNITLTHC KGDSCNGDNL  
       151 LDEEAPSKSE FENLINESERI EKYKKDGKSD KFTNLVATAV QANGTNKYVI  
       201 IYKDKSASSS SARFRRSARS RRSLPAEMPL IPVNQADTLI VDGEAVSLTG  
       251 HSGNIFAPEG NYRYLTYGAE KLPGGSYALR VQGEPAKgem LAGTAVYN  
       301 VLHFHTENG PYPTRGRFAA KVDFGSKSVD GIIDSQDDLH MGTQKFKA  
       351 DGNGFKGTWT ENGGGDVSGR FYGPAGEEEVA GKYSYRPTDA EKGFFGVFAG  
       401 KKEQDGSGGG GATYKVDEYH ANARFAIDHF NTSTNVGGFY GLTGSVEFDQ  
       451 AKRDGKIDIT IPVANLQSGS QHFTDHLSA DIFDAQYPD IRFVSTKFNF  
       501 NGKKLVSVDG NLTMHGKTAP VKLKAEKFNC YQSPMAKTEV CGGDFSTTID  
       551 RTKWDVSYLV NVGMTKSVRI DIQIEAAKQ\*

ΔG287-961

5	1	ATGGCTAGCC CCGATGTTAA ATCGGCGGAC ACGCTGTCAA AACCGGCCGC
	51	TCCCTGTTGTT GCTGAAAAAG AGACAGAGGT AAAAGAAGAT GCGCCACAGG
	101	CAGGTTCTCA AGGACAGGGC CGGCCATCCA CACAAGGCAG CCAAGATATG
	151	GCGGCAGTTT CGGCAGAAA TACAGGCAAT GGCGGTGCGG CAACAACGGA
	201	CAAACCCAAA AATGAAGACG AGGGACCGCA AAATGATATG CCGCAAAATT
	251	CCGCCGAATC CGCAATCAA ACAGGAAACA ACCAACCCGC CGATTCTTC
	301	GATTCCGCCCG CCGCGTCAA CCCTGCACCT GCAGATGGCG GTAGCAATT
	351	TGGAAGGGTT GATTGGCTA ATGGCGTTT GATTGATGGG CGCTCGCAA
10	401	ATATAACGTT GACCCACTGT AAAGGCGATT CTTGTAATGG TGATAATT
	451	TTGGATGAAG AAGCACCGTC AAAATCAGAA TTGAAAATT TAAATGAGTC
	501	TGAACGAATT GAGAAATATA AGAAAAGATGG GAAAAGCGAT AAATTTACTA
	551	ATTTGGTTGC GACAGCAGTT CAAGCTAATG GAACTAACAA ATATGTCATC
	601	ATTTATAAAG ACAAGTCCGC TTCATCTTC TCTGCGCGAT TCAGGCGTT
	651	TGCACGGTCG AGGAGGTCGC TTCCCTGCCGA GATGCCGCTA ATCCCCGTCA
15	701	ATCAGGCGGA TACGCTGATT GTCGATGGGG AAGCGCGAT CCTGACGGGG
	751	CATTCCGGCA ATATCTCGC GCCCAGGGG AATTACCGGT ATCTGACTTA
	801	CGGGGCGGAA AAATTGCGGC GCGGATCGTA TGCCCTCCGT GTGCAAGGCG
	851	AACCGGCAAA AGGCGAAATG CTTGCTGGCA CGGCGTGTAA CACGGCGAA
	901	GTGCTGCATT TTCATACGGG AAACGGCGT CCGTACCCGA CTAGAGGCAG
	951	GTGGCCGCA AAAGTCGATT TCGGCAGCAA ATCTGTTGGAC GGCAATTATCG
20	1001	ACAGCGGCGA TGATTTGCAT ATGGGTACGC AAAAATTCAA AGCCGCCATC
	1051	GATGGAAACG GCTTTAAGGG GACTTGGACG GAAAATGGCG GCGGGGATGT
	1101	TTCCCGGAAGG TTTTACGGCC CGGCCGGCGA GGAAGTGGCG GAAAATACAA
	1151	GCTATCGCCC GACAGATGCG GAAAAGGGCG GATTCCCGT GTTTGCCGGC
	1201	AAAAAAGAGC AGGATGGATC CGGAGGAGGA GGAGCCACAA ACCACGACGA
	1251	TGTTAAAAAA GCTGCCACTG TGGCCATTGC TGCTGCCCTAC ACAATGGCC
25	1301	AAGAAATCAA CGGTTTCAA GCTGGAGAGA CCATCTACGA CATTGATGAA
	1351	GACGGCACAA TTACAAAAAA AGACGCAACT GCAGCCGATG TTGAAGCCGA
	1401	CGACTTTAAA GGTCTGGGT TGAAAAAAAGT CGTACTAAC CTGACCAAAA
	1451	CCGTCAATGA AAACAAACAA AACGTCGATG CCAAAGTAAA AGCTGCAGAA
	1501	TCTGAAATAG AAAAGTTAAC AACCAAGTTA GCAGACACTG ATGCCGTTT
30	1551	AGCAGATACT GATGCCGCTC TGGATGCAAC CACCAACGCC TTGAATAAT

35	1601	TGGGAGAAAA TATAACGACA TTTGCTGAAG AGACTAAGAC AAATATCGTA
	1651	AAAATTGATG AAAAATTAGA AGCCGTGGCT GATACCGTCG ACAAGCATGC
	1701	CGAACGCATTC AACGATATCG CCGATTCTATT GGATGAAACC AACACTAAGG
	1751	CAGACGAAGC CGTCAAAACC CGCAATGAAG CCAAACAGAC GGCGAAGAA
	1801	ACCAAACAAA ACGTCGATGC CAAAGTAAA GCTGCAGAAA CTGCAGCAGG
	1851	CAAAGCCGAA GCTGCCGCTG GCACAGCTAA TACTGCAGCC GACAAGGCCG
	1901	AAGCTGTCGC TGCAAAAGTT ACCGACATCA AAGCTGATAT CGCTACGAAC
	1951	AAAGATAATA TTGCTAAAAA AGCAAAACAGT GCCGACGTGT ACACCAGAGA
40	2001	AGAGTCTGAC AGCAAAATTG TCAGAATTGA TGGCTGAAAC GCTACTACCG
	2051	AAAAATTGGA CACACGCTTG GCTTCTGCTG AAAAATCCAT TGCGATCAC
	2101	GATACTCGCC TGAACGGTTT GGATAAAAACA GTGTCAGACC TCGCAAAGA
	2151	AACCCGCCAA GGCCTTGCAG AACAGCCGC GCTCTCCGGT CTGTTCCAAC
	2201	CTTACAAACGT GGGTCGGTTC AATGTAACGG CTGCAGTCGG CGGCTACAAA
	2251	TCCGAATCGG CAGTCGCCAT CGGTACCGGC TTCCGTTTA CGAAAACCTT
	2301	TGCCGCCAAA GCAGGCCTGG CAGTCGGCAC TTGTCGGGT TCTTCCGCAG
45	2351	CCTACCATGT CGGCGTCAAT TACGAGTGGT AACTCGAG

1 MASPDVKSAAD TLSKPAAPVV AEKETEVKED APQAGSQGQG APSTQGSQDM  
 5 51 AAVSAENTGN GGAATTDKPK NEDEGPQNDM PQNSAESANQ TGNNQPADSS  
 101 DSAPASNPAP ANGGSNFGRV DLANGVLIDG PSQNITLTHC KGDSCNGDNL  
 151 LDEEAPSKE FENLNESERI EKYKKDGKSD KFTNLVATAV QANGTNKYVI  
 201 IYKDKSASSS SARFRRSARS RRSLPAEMPL IPVNQADTLI VDGEAVSLTG  
 251 HSGNIFAPEG NYRYLTYGAE KLPGGSYALR VQGEPAKgem LAGTAVYNGE  
 301 VLHFHTENGR PYPTGRFAA KVDFGSKSVD GIIDSGDDLH MGTQKFKAII  
 351 DGNGFKGTWT ENGGGDVSGR FYGPAGEEVA GKYSYRPTDA EKGGFGVFAG  
 401 KKEQDGSGGG GATNDDDVKK AATVAIAAAY NNGQEINGFK AGETIYDIDE  
 451 DGTITKKDAT AADVEADDK GLGLKKVVTN LTKTVNENKQ NVDAKVAAE  
 501 SEIEKLTTKL ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV  
 551 KIDEKLEAVA DTVDKHAEAF NDIADSLSDET NTKADEAVKT ANEAQQTAE  
 601 TKQNVDAKVK AAETAAGKAE AAAGTANTAA DKAEEAVAALK TDIKADIATN  
 651 KDNIAKKANS ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEEKSIADH  
 701 DTRLNGLDKT VSDLRKETRQ GLAEQAALSG LFQPYNVGRF NVTAAVGGYK  
 751 SESAVAIGTG FRFTENFAAK AGVAVGTSSG SSAAYHVGVN YEW\*

	<b>ELISA</b>	<b>Bactericidal</b>
20	ΔG287-953-His	3834 65536
	ΔG287-961-His	108627 65536

25 [0219] The bactericidal efficacy (homologous strain) of antibodies raised against the hybrid proteins was compared with antibodies raised against simple mixtures of the component antigens (using 287-GST) for 919 and ORF46.1:

	<b>Mixture with 287</b>	<b>Hybrid with ΔG287</b>
30	919 32000	128000
	ORF46.1 128	16000

35 [0220] Data for bactericidal activity against heterologous MenB strains and against serotypes A and C were also obtained:

	<b>919</b>		<b>ORF46.1</b>		
	<i>Strain</i>	<i>Mixture</i>	<i>Hybrid</i>	<i>Mixture</i>	<i>Hybrid</i>
40	<b>NGH38</b>	1024	32000	-	16384
	<b>MC58</b>	512	8192	-	512
45	<b>BZ232</b>	512	512	-	-
	<b>MenA (F6124)</b>	512	32000	-	8192
	<b>MenC (C11)</b>	>2048	>2048	-	-
	<b>MenC (BZ133)</b>	>4096	64000	-	8192

50 [0221] The hybrid proteins with ΔG287 at the N-terminus are therefore immunologically superior to simple mixtures, with ΔG287-ORF46.1 being particularly effective, even against heterologous strains. ΔG287-ORF46.1K may be expressed in pET-24b.

[0222] The same hybrid proteins were made using New Zealand strain 394/98 rather than 2996:

AG287NZ-919

5	1	ATGGCTAGCC	CCGATGTCAA	GTCGGCGGAC	ACGCTGTCAA	AACCTGCCGC
	51	CCCTGTTGTT	TCTGAAAAAAG	AGACAGAGGC	AAAGGAAGAT	GCGCCACAGG
	101	CAGGTTCTCA	AGGACAGGGC	GCGCCATCCG	CACAAGGCGG	TCAAGATATG
	151	GCGGCGGTTT	CGGAAGAAAA	TACAGGCAAT	GCGGTCGGG	CAGCAACGGA
	201	CAAACCCAAA	AATGAAGACC	AGGGGCGCA	AAATGATATG	CCGCAAAATG
	251	CCGCGATAC	AGATAGTTG	ACACCGAATC	ACACCCGGC	TCGAATATG
	301	CCGGCCGGAA	ATATGGAAAA	CCAAGCACCG	GATGCCGGGG	AATCGGAGCA
	351	GCGGGCAAC	CAACCGGATA	TGGCAAATAC	GGCGGACGGA	ATGCAGGGTG
10	401	ACGATCCGTC	GGCAGGGCGGG	GAAAATGCCG	GCAATACGGC	TGCCCAAGGT
	451	ACAAATCAAG	CCGAAAACAA	TCAAACCGCC	GGTTCTCAA	ATCCTGCCCTC
	501	TTCAACCAAT	CCTAGCGCCA	CGAATAGCGG	TGGTGATTTT	GGAAGGACGA
	551	ACGTGGGCAA	TTCTGTTGTC	ATTGACGGGC	CGTCGCAAA	TATAACGTTG
	601	ACCCACTGTA	AAGGCGATTC	TTGTAGTGGC	AATAATTCT	TGGATGAAGA
15	651	AGTACAGCTA	AAATCAGAAT	TTGAAAAATT	AAGTGTGCA	GACAAAATAA
	701	GTAATTACAA	GAAAATGGG	AAGAATGACG	GGAAGAATGA	TAATTTGTC
	751	GTTTTGGTT	CCGATAGTGT	GCAGATGAAG	GGAATCAATC	AATATATTAT
	801	CTTTTATAAA	CCTAAACCCA	CTTCATTTCG	GCGATTAGG	CGTTCTGCAC
	851	GGTCGAGGGC	GTCGCTTCGG	GCCGAGATGC	CGCTGATTCC	CGTCAATCAG
	901	GCGGATACGC	TGATTGTCGA	TGGGAAGCG	GTCAGCCTGA	CGGGGCATTC
20	951	CGGCAATATC	TTCGCGCCCG	AAGGGAATTA	CCGGTATCTG	ACTTACGGGG
	1001	CGGAAAATT	GCCCCGGCGA	TCGTATGCC	TCCGTGTTCA	AGGCAGAACCT
	1051	TCAAAAGGCG	AAATGCTCGC	GGGCACGCCA	GTGTACAACG	CGGAAGTGCT
	1101	GCATTTTCAT	ACGGAAAACG	GCCGTCGTC	CCCGTCCAGA	GGCAGGTTTG
	1151	CCGCAAAGT	CGATTTCGGC	AGCAAATCTG	TGGACGGCAT	TATCGACAGC
	1201	GGCGATGGTT	TGCATATGGG	TACGCAAAA	TTCAAAGCCG	CCATCGATGG
25	1251	AAACGGCTT	AAGGGGACTT	GGACGGAAA	TGGCGCCGGG	GATGTTTCCG
	1301	GAAAGTTTA	CGGCCCCGGC	GGCGAGGAAG	TGGCGGGAAA	ATACAGCTAT
	1351	CGCCCAACAG	ATGCGGAAAA	GGGCGGATT	GGCGTGTGTTG	CCGGCAAAAAA
	1401	AGAGCAGGAT	GGATCCGGAG	GAGGAGGATG	CCAAAGCAAG	AGCATCCAAA
	1451	CCTTCCGCA	ACCCGACACA	TCCGTATCA	ACGGCCCGGA	CCGGCCGGTC
	1501	GGCATCCCCG	ACCCGCCGG	AACGACGGTC	GGCGGCGGCG	GGGCCGTCTA
30	1551	TACCGTTGTA	CCGCACCTGT	CCCTGCCCA	CTGGGCGGCG	CAGGATTTCG
	1601	CCAAAAGCCT	GCAATCCTTC	CGCCTCGGCT	GCGCAATT	GAAAACCGC
	1651	CAAGGCTGGC	AGGATGTGTC	CGCCCAAGCC	TTTCAAACCC	CCGTCCATT
	1701	CTTCAGGC	AAACAGTTT	TTGAACGCTA	TTTCACGCCG	TGGCAGGTTG
	1751	CAGGCAACGG	AAGCCTTGCC	GGTACGGTTA	CCGGCTATT	CGAGCCGGTG
	1801	CTGAAGGGCG	ACGACAGGGC	GACGGCACAA	GCCGCTTCC	CGATTTACGG
35	1851	TATTCGGAC	GATTTTATCT	CCGTCCCCCT	GCCTGCCGGT	TTGCAGGAGCG
	1901	GAAAAGCCT	TGTCCGCATC	AGGCAGACGG	GAAAAAACAG	CGGCACAATC
	1951	GACAATACCG	CGGGCACACACA	TACCGCCGAC	CTCTCCGAT	TCCCCATCAC
	2001	CGCGCGACA	ACGGCAATCA	AAGGCAGGTT	TGAAGGAAGC	CGCTTCCCTCC
40	2051	CCTACCACAC	GCGCAACCAA	ATCAACGGCG	GCGCGCTTGA	CGGCAAAGCC
	2101	CCGATACTCG	GTTACGCCGA	AGACCCCGTC	GAACTTTTT	TTATGCACAT
	2151	CCAAGGCTCG	GGCGCTCTGA	AAACCCCGTC	CGGCAAATAC	ATCCGCATCG
	2201	GCTATGCCGA	AAAAAACGAA	CATCCCTACG	TTTCCATCGG	ACGCTATATG
	2251	GCGGACAAAG	GCTACCTCAA	GCTCGGGCAG	ACCTCGATGC	AGGGCATCAA
45	2301	AGCCTATATG	CGGCAAAATC	CGCAACGCC	CGCCGAAGTT	TTGGGTCAA
	2351	ACCCCAGCTA	TATCTTTTC	CGCGAGCTTG	CCGGAAGCAG	CAATGACGGT
	2401	CCCCTCGGGC	CACTGGGCAC	GCCGTTGATG	GGGAATATG	CCGGCGCAGT
	2451	CGACCGGCAC	TACATTACCT	TGGGCGGCC	CTTATTGTC	GCCACCGCCC
	2501	ATCCGGTTAC	CCGCAAAGCC	CTCAACCGCC	TGATTATGGC	GCAGGATACC
50	2551	GGCAGCGCGA	TTAAAGGC	GGTGCAGCGT	GATTATTTT	GGGGATACGG
	2601	CGACGAAGCC	GGCGAACTTG	CCGGCAAACCA	GAAAACCACG	GGTTACGTCT
	2651	GGCAGCTCCT	ACCCAACGGT	ATGAAGCCCC	AATACCGCCC	GTAAAAGCTT

1 MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQGQG APSAQGGQDM  
 51 AAVSEENTGN GGAAATDKPK NEDEGAQNMD PQNAADTDSL TPNHTPASN  
 101 PAGNMENQAP DAGESEQPAN QPMANTADG MQGDDPSAGG ENAGNTAAQG  
 151 TNQAENNQTA GSQNMPASSTN PSATNSGGDF GRTNVGNSVV IDGPSQNITL  
 201 THCKGDSCSG NNFLDEEVQL KSEFEKLSDA DKISNYKKDG KNDGKNDKFV  
 251 GLVADSVQMK GINQYIIFYK PKPTSFARFR RSARSRRSLP AEMPLIPVNQ  
 301 ADTLIVDGEA VSLTGHSIGNI FAPEGNYRYL TYGAEKLPGG SYALRVQGE  
 351 SKGEMLAGTA VYNGEVLHFFH TENGRPSPL GRFAAKVDFG SKSVDGIIDS  
 401 GDGLHMGTK FKAADIDNGF KGTWTENGGA DVSGKFYGP AEEVAGKYSY  
 451 RPTDAEKGGF GVFAKGKEQD GSggggcosk SIQTFQPDT SVINGPDRPV  
 501 GIPDPAGTTV GGGGAVYTVV PHLSLPHWAA QDFAKSLQSF RLGCANLKNR  
 551 QGWQDVCAQA FQTPVHSFQA KQFFERYFTP WQVAGNGSLA GTVTGYYEPV  
 601 LKGDDRRTAQ ARFPYIGIPD DFISVPLPAG LRSGKALVRI RQTGKNSGTI  
 651 DNTGGHTAD LSRFPIART TAIKGRFEGS RFLPYHTRNQ INGGALDGKA  
 701 PILGYAEDPV ELFFMHIQGS GRLKTPSGKY IRIGYADKNE HPYVSIGRYM  
 751 ADKGYLKLQG TSMQGIKAYM RQNQPQLAEV LGQNPSYIFF RELAGSSNDG  
 801 PVGALGTPLM GEYAGAVDRH YITLGAPLFV ATAHPVTRKA LNRLIMAQDT  
 851 GSAIKGAVRV DYFWGYGDEA GELAGKQKTT GYVWQLLPNG MKPEYRP\*

20

**ΔG287NZ-953**

1 ATGGCTAGCC CCGATGTCAA GTCGGCGGAC ACGCTGTCAA AACCTGCCGC  
 51 CCCTGTTGTT TCTGAAAAAG AGACAGAGGC AAAGGAAGAT GCGCCACAGG  
 101 CAGGTTCTCA AGGACAGGGC GCGCCATCCG CACAAGGCGG TCAAGATATG  
 151 GCGGCGGTTT CGGAAGAAAA TACAGGCAAT GGCCTGCGG CAGCAACGGA  
 201 CAAACCCAAA AATGAAGACG AGGGGGCGCA AAATGATATG CCGCAAAATG  
 251 CCGCCGATAC AGATAGTTG ACACCGAATC ACACCCCGGC TTGCAATATG  
 301 CCGCCCGGAA ATATGGAAA CCAAGCACC GATGCCGGG AATCGGAGCA  
 351 GCCGGCAAAC CAACCGATA TGGCAAATAC GGCGGACGGA ATGCAGGGTG  
 401 ACGATCGTC GGCAGGCGGG GAAATGCGC GCAATACGGC TGCCCAAGGT  
 451 ACAAAATCAAG CCGAAAACAA TCAAAACGCC GGTTCTCAAA ATCCTGCC  
 501 TTCAACCAAT CCTAGCGCCA CGAATAGCGG TGGTGATTTC GGAAGGACGA  
 551 ACGTGGGCAA TTCTGTTGTG ATTGACGGGC CGTCGCAAA TATAACGTTG  
 601 ACCCACTGTA AAGGCAGATT TTGTAGTGGC AATAATTCT TGGATGAAGA  
 651 AGTACAGCTA AAATCAGAAT TTGAAAATT AAGTGATGCA GACAAAATAA  
 701 GTAATTACAA GAAAGATGGG AAGAATGACG GGAAGAATGA TAAATTTGTC  
 751 GGTTTGGTTG CCGATAGTGT GCAGATGAAG GGAATCAATC AATATATTAT  
 801 CTTTTATAAA CCTAACCCCA CTTCATTTGC GCGATTAGG CGTCTGCAC  
 851 GGTCGAGGCG GTCGCTTCCG CGCGAGATGC CGCTGATTCC CGTCAATCAG  
 901 GCGGATACGC TGATTGTCGA TGGGGAAGCG GTCAGCCTGA CGGGGCATTC  
 951 CGGCAATATC TTCGCGCCCG AAGGGAATTAA CGGGTATCTG ACTTACGGGG  
 1001 CGGAAAATT GCCCCGGCGGA TCGTATGCC TCCGTGTTCA AGGCGAACCT  
 1051 TCAAAAGCG AAATGCTCGC GGGCACGGCA GTGTACAACG GCGAAGTGC  
 1101 GCATTTTCAT ACGGAAAACG GCGTCCGTC CCCGTCAGA GGCAGGTTG  
 1151 CCGAAAAGT CGATTCGGC AGCAAATCTG TGGACGGCAT TATCGACAGC  
 1201 GGCGATGGTT TGCATATGGG TAGCAAAAA TTCAAAGCCG CCATCGATGG  
 1251 AAACGGCTTT AAGGGGACTT GGACGGAAA TGGCGCGGG GATGTTCCG  
 1301 GAAAGTTTA CGGCCCGGCC GGCAGGAAAG TGGCGGGAAA ATACAGCTAT  
 1351 CGCCCAACAG ATGCGGAAAA GGGCGGATTG GGCCTGTTG CCGGCAAAAA  
 1401 AGAGCAGGAT GGATCCGGAG GAGGAGGAGC CACCTACAAA GTGGACGAAT  
 1451 ATCACGCCAA CGCCCGTTTC GCCATCGACC ATTTCAACAC CAGCACCAAC  
 1501 GTCGGCGGTT TTACCGGTCT GACCGGTTCC GTCGAGTTCG ACCAAGCAAA  
 1551 ACGCGACGGT AAAATCGACA TCACCATCCC CGTTGCCAAC CTGCAAAGCG

50

55

**EP 1 790 660 A2**

1601 GTTCGCAACA CTTTACCGAC CACCTGAAAT CAGCCGACAT CTTCGATGCC  
1651 GCCCAATATC CGGACATCCG CTTTGTTC ACCAAATTCA ACTTCAACGG  
1701 CAAAAAACTG GTTTCGTTG ACGGCAACCT GACCATGCAC GGCAAAACCG  
1751 CCCCCGTCAA ACTCAAAGCC GAAAAATTCA ACTGCTACCA AAGCCCAGTG  
1801 GCGAAAACCG AAGTTTGCAG CGCGACTTC AGCACCA TCGACCGCAC  
1851 CAAATGGGC GTGGACTACC TCGTTAACGT TGGTATGACC AAAAGCGTCC  
1901 GCATCGACAT CCAAATCGAG GCAGCCAAAC AATAAAAGCT T

10 1 MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQGQG APSAQGGQDM  
51 AAVSEENTGN GGAAATDKPK NEDEGAQNMD PQNAADTDLSL TPNTHTPASNM  
101 PAGNMENQAP DAGESEQPAN QPDMANTADG MQGDDPSAGG ENAGNTAAQG  
151 TNQAENNQTA GSQNPARSTN PSATNSGGDF GRTNVGNSVV IDGPSQNITL  
201 THCKGDSCSG NNFLDEEVQL KSEFEKLSDA DKISNYKKDG KNDGKNDKFV  
251 GLVADSVQMK GINQYIIFYK PKPTSFARFR RSARSSRSLP AEMPLIPVNQ  
301 ADTLIVDGEA VSLTGHSGNI FAPEGNYRYL TYGAEKLPGG SYALRVQGEF  
351 SKGEMLAGTA VYNGEVLHFH TENGRPSPSR GRFAAKVDFG SKSVDGIIDS  
401 GDGLHMGTK FKAAIDGNF KGTWTENGGS DVSGKFYGP AEEVAGKYSY  
451 RPTDAEKGGF GVFAGKKEQD GSAGGGATYK VDEYHANARF AIDHFNTSTN  
501 VGGFYGLTGS VEFDQAKRDG KIDITIPVAN LQSGSQHFTD HLKSADIFDA  
551 AQYPDIFVS TKFNFNGKKL VSVDGNLTMH GKTAPVKLKA EKFNCYQSPM  
601 AKTEVCGGDF STTIDRTKWG VDYLNVGMT KSVRIDIQIE AAKQ\*

25

30

35

40

45

50

55

ΔG287NZ-961

5	1	ATGGCTAGCC CCGATGTCAA GTCGGCGGAC ACGCTGTCAA AACCTGCCGC
	51	CCCTGTTGTT TCTGAAAAAG AGACAGAGGC AAAGGAAGAT GCGCCACAGG
	101	CAGGTTCTCA AGGACAGGGC GCGCCATCG CACAAGGC GGCGGTGCGG CAGCAACCGA
	151	GCGGCGGTTT CGGAAGAAAA TACAGGCAAT GGCGGTGCGG CAGCAACCGA
	201	CAAACCCAAA AATGAAGACG AGGGGGCGCA AAATGATATG CCGCAAAATG
	251	CCGCGATAC AGATAGTTG ACACCGAAC ACACCCGGC TTGAAATATG
10	301	CCGGCCGGAA ATATGGAAAA CCAAGCACCG GATGCCGGGG AATCGGAGCA
	351	GCGGCAAAAC CAACCGATA TGGCAAATAC GGCGGACGGA ATGCAGGGTG
	401	ACGATCCGTC GGCAGGCGGG GAAAATGCG GCAATACGGC TGCCCAAGGT
	451	ACAAATCAAG CCGAAAACAA TCAAACCGCC GGTTCTCAA ATCCTGCCTC
	501	TTCAACCAAT CCTAGCGCCA CGAATAGCGG TGGTAGTTT GGAAGGACGA
	551	ACGTGGCAA TTCTGTTGTT ATTGACGGGC CGTCGCAAAA TATAACGTTG
15	601	ACCCACTGTA AAGCGATTG TTGTAATGCG AATAATTCT TGATGAAGA
	651	AGTACAGCTA AAATCAGAAT TTGAAAATT AAGTGTGCA GACAAAATAA
	701	GTAATTACAA GAAAGATGGG AAGAATGACG GGAAGAATGA TAAATTGTC
	751	GGTTTGGTTG CCGATAGTGT GCAGATGAAG GGAATCAATC AATATATTAT
	801	CTTTTATAAA CCTAAACCCA CTTCATTTGC GCGATTAGG CGTTCTGCAC
	851	GGTCGAGGCG GTCGCTTCGGC GCCGAGATGC CGCTGATTCC CGTCAATCAG
20	901	GCGGATACGC TGATTGTCGA TGGGGAAAGCG GTCAGCCTGA CGGGGCATTC
	951	CGGCAATATC TTCGCGCCCG AAGGGAAATTA CCGGTATCTG ACTTACGGGG
	1001	CGGAAAATT GCCCCGGCGA TCGTATGCC TCCGTGTTCA AGGCACCT
	1051	TCAAAAGGCG AAATGCTCGC GGGCACGGCA GTGTACAACG GCGAAGTGCT
	1101	GCATTTTCAT ACGAAAACAG GCGTCCCGC CCCGTCAGA GGCAGGTTG
	1151	CCGCAAAAGT CGATTCGGC AGCAAATCTG TGGACGGCAT TATCGACAGC
	1201	GGCGATGGTT TGCAATATGGG TACGCAAAA TTCAAAGCCG CCATCGATGG
25	1251	AAACGGCTT AAGGGGACTT GGACGGAAA TGGCGCGGG GATGTTCCG
	1301	GAAAGTTTA CGGCCCGGCC GCGAGGAAG TGGCGGGAAA ATACAGCTAT
	1351	CGCCCAACAG ATGCGGAAAA GGGCGGATTC GCGTGTGTTG CCGGCAAAAA
	1401	AGAGCAGGAT GGATCCGGAG GAGGAGGAGC CACAAACGAC GACGATGTTA
	1451	AAAAAGCTGC CACTGTGGCC ATTGCTGCTG CCTACAACAA TGGCCAAGAA
30	1501	ATCAACGTT TCAAAGCTGG AGAGACCAC TACGACATTG ATGAAGACGG
	1551	CACAATTACCA AAAAAGACG CAACTGCAGC CGATGTTGAA GCGGACGACT
	1601	TTAAAGGTCT GGGTCTGAA AAAGTCGTGA CTAACCTGAC CAAAACGTC
	1651	AATGAAAACA AACAAACACGT CGATGCCAA GTAAAGCTG CAGAATCTGA
	1701	AATAGAAAAG TAAACACCA AGTTAGCAGA CACTGATGCC GCTTTAGCAG
	1751	ATACTGATGC CGCTCTGGAT GCAACCACCA ACGCCTTGAA TAAATTGGGA
35	1801	GAAAATATAA CGACATTG TGAAAGAGACT AAGACAAATA TCGTAAAAT
	1851	TGATGAAAAA TTAGAAGCCG TGGCTGATAC CGTCGACAAG CATGCCGAAG
	1901	CATTCAACGA TATCGCCGAT TCATTGGATG AAACCAACAC TAAGGCAGAC
	1951	GAAGCCGTC AAACGCCAA TGAAGCCAA CAGACGGCG AAGAAACCAA
	2001	ACAAAACGTC GATGCCAAAG TAAAAGCTGC AGAAACTGCA GCAGGCCAAAG
	2051	CCGAAGCTGC CGCTGGCACA GCTAATACTG CAGCCGACAA GGCGGAAGCT
40	2101	GTCGCTGCAA AAGTTACCGA CATCAAAGCT GATATCGCTA CGAACAAAGA

45	2151	TAATATTGCT AAAAAGCAA ACAGTGGCGA CGTGTACACC AGAGAAGAGT
	2201	CTGACAGCAA ATTTGTCAGA ATTGATGGTC TGAACGCTAC TACCGAAAAA
	2251	TTGGACACAC GCTTGGCTTC TGCTGAAAAA TCCATTGCCG ATCACGATAC
	2301	TCGCCTGAAC GGTGGATA AAACAGTGT AGACCTGCGC AAAGAAACCC
	2351	GCCAAGGCCT TGCAGAACAA GCGCGCTCT CGGCTGTGTT CCAACCTTAC
50	2401	AACGTGGTC GGTTCAATGT AACGGCTGCA GTCGGCGGCT ACAAAATCCGA
	2451	ATCGGCAGTC GCCATCGGTA CCGGCTTCCG CTTTACCGAA AACTTGCCG
	2501	CCAAAGCAGG CGTGGCAGTC GGCACCTCGT CGGTTCTTC CGCAGCCTAC
	2551	CATGTCGGCG TCAATTACGA GTGGTAAAAG CTT

1 MASPDVKSAD TLSKPAAPVV SEKETEAKED APQAGSQGQG APSAQGGQDM  
 51 AAVSEENTGN GGAAATDKPK NEDEGAQNMD PQNAADTDSL TPNHTPASNMR  
 101 PAGNMENQAP DAGESEQPAN QPDMANTADG MQGDDPSAGG ENAGNTAAQG  
 151 TNQAENNQTA GSQNPPASSTN PSATNSGGDF GRTNVGNSVV IDGPSQNITL  
 201 THCKGDSCSG NNFLDEEVQL KSEFEKLSDA DKISNYKKDG KNDGKNDKFV  
 251 GLVADSVQMK GINQYIIFYK PKPTSFARFR RSARSRRSLP AEMPLIPVNQ  
 301 ADTLIVDGEA VSLTGHSGNI FAPEGNYRYL TYGAEKLPGG SYALRVQGEPE  
 351 SKGEMLAGTA VYNGEVLHFH TEGRPSPSR GRFAAKVDFG SKSVDGIIDS  
 401 GDGLHMGTK FKAAIDGNF KGTWTENGGS DVSGKFYGPA GEEVAGKYSY  
 451 RPTDAEKGGF GVFAKGKEQD GSGGGGATND DDVKKAATVA IAAAYNNNGQE  
 501 INGFKAGETI YDIDEDEGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV  
 551 NENKQNVDAK VKAEESEIEK LTTKLADTDALADTAALD ATTNALNKLG  
 601 ENITTFAEET KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNTKAD  
 651 EAVKTANEAK QTAEETKQNV DAKVKAETA AGKAEAAAAGT ANTAADKAEA  
 701 VAAKVTDIKA DIATNKDNIA KKANSADVYT REESDSKFVR IDGLNATTEK  
 751 LDTRLASAEK SIADHDTRLN GLDKTVSDLR KETRQGLAEQ AALSGLFQPY  
 801 NVGRFNVTAA VGGYKSESACV AIGTGFRFTE NFAAKAGVAV GTSSGSSAAY  
 851 HVGVNYEW\*

20 *ΔG983 and hybrids*

[0223] Bactericidal titres generated in response to ΔG983 (His-fusion) were measured against various strains, including the homologous 2996 strain:

25		2996	NGH38	BZ133
	ΔG983	512	128	128

[0224] ΔG983 was also expressed as a hybrid, with ORF46.1, 741, 961 or 961c at its C-terminus:

30 **ΔG983-ORF46.1**

1	ATGACTTCTG CGCCCGACTT CAATGCAGGC GGTACCGGTA TCGGCAGCAA
51	CAGCAGAGCA ACAACAGCGA AATCAGCAGC AGTATCTTAC GCCGGTATCA
101	AGAACGAAAT GTGCAAAGAC AGAAGCATGC TCTGTGCCGG TCGGGATGAC
151	GTTGCGGTTA CAGACAGGGG TGCCAAAATC AATGCCCCCCC CCCCAGAATCT
201	GCATACCGGA GACTTTCCAA ACCCAAATGA CGCATAACAAG AATTGATCA
251	ACCTCAAAAC TGCAATTGAA GCAGGCTATA CAGGACGCGG GGTAGAGGTA
301	GGTATCGTCG ACACAGGCAG ATCCGTCGGC AGCATATCCT TTCCCGAACT
351	GTTGGCAGA AAAGAACACG GCTATAACGA AAATTACAAA AACTATAACGG
401	CGTATATGCG GAAGGAAGCG CCTGAAGACG GAGGCGGTAA AGACATTGAA
451	GCTTCTTTCG ACGATGAGGC CGTTATAGAG ACTGAAGCAA AGCCGACGGG
501	TATCCGCCAC GTAAAAGAAA TCGGACACAT CGATTGGTC TCCCATATTA
551	TTGGCGGGCG TTCCGTGGAC GGCAGACCTG CAGGGCGTAT TCGGCCCGAT
601	GCGACGCTAC ACATAATGAA TACGAATGAT GAAACCAAGA ACGAAATGAT
651	GGTTGCAGCC ATCCCAATG CATGGGTCAA GCTGGGCAGA CGTGGCGTGC
701	GCATCGTCAA TAACAGTTT GGAACAACAT CGAGGGCAGG CACTGCCGAC
751	CTTTTCCAAA TAGCCAATTC GGAGGAGCAG TACCGCAAG CGTTGCTCGA
801	CTATTCCGGC GGTGATAAAAA CAGACGAGGG TATCCGCCTG ATGCAACAGA
851	GCGATTACGG CAACCTGTCC TACCACATCC GTAATAAAA CATGTTTTC
901	ATCTTTTCGA CAGGCAATGA CGCACAAAGCT CAGCCCAACA CATATGCCCT
951	ATTGCCATT TATGAAAAAG ACGCTCAAAA AGGCATTATC ACAGTCGCAG
1001	GCGTAGACCG CAGTGGAGAA AAGTTCAAAC GGGAAATGTA TGGAGAACCG
1051	GGTACAGAAC CGCTTGAGTA TGGCTCCAAC CATTGCGGAA TTACTGCCAT
1101	GTGGTGCTG TCGGCACCCCT ATGAAGCAAG CGTCCGTTTC ACCCGTACAA

5	1151	ACCCGATTCA AATTGCCGGA ACATCCTTTT CCGCACCCAT CGTAACCGGC
	1201	ACGGCGGCTC TGCTGCTGCA GAAATACCCG TGGATGAGCA ACGACAACCT
	1251	GCGTACCAACG TTGCTGACCGA CCGCTCAGGA CATCGGTGCA GTCGGCGTGG
	1301	ACAGCAAGTT CGGCTGGGA CTGCTGGATG CGGGTAAGGC CATGAACGGGA
	1351	CCCGCGTCCT TTCCGTTCGG CGACTTTACGCCGATAACGA AAGGTACATC
	1401	CGATATTGCC TACTCCTTCC GTAACGACAT TTCAGGCACG GGCGGCCTGA
	1451	TCAAAAAGG CGGCAGCCAA CTGCAACTGC ACGGCAACAA CACCTATACG
	1501	GGCAAAACCA TTATCGAAGG CGGTTCGCTG GTGTTGTACG GCAACAACAA
	1551	ATCGGATATG CGGCTCGAAA CAAAGGTCG GCTGATTAT AACGGGGCGG
	1601	CATCCCGCGG CAGCGTGAAC AGCGACGGCA TTGCTATCT GGCAGATACC
10	1651	GACCAATCCG CGCGAACCGA ACCCGTACAC ATCAAAGGCA GTCTCGAGCT
	1701	GGACGGAAA GGTACCGCTGT ACACACGTTT GGGCAAATCTG CTGAAAGTGG
	1751	ACGGTACGGC GATTATCGGC GGCAAGCTGT ACATGTCGGC ACGGCGCAAG
	1801	GGGGCAGGCT ATCTCAACAG TACCGGACGA CGTGTCCCT TCCTGAGTGC
	1851	CGCCAAAATC GGGCAGGATT ATTCTTTCTT CACAAACATC GAAACCGACG
	1901	GC GGCGTCG GCGTCCCTC GACAGCGTCG AAAAACACAG GGGCAGTGAA
	1951	GGCGACACGC TGTCCTATT TGTCCTCGC GGCAATGCGG CACGGACTGC
	2001	TTCCGGCAGC GCACATTCCG CGCCCGCCGG TCTGAAACAC GCGTAGAAC
	2051	AGGGCGGCAG CAATCTGGAA AACCTGATGG TCGAACTTGG TGCCCTCGAA
	2101	TCATCCGCAA CACCCGAGAC GGTTGAAACT GCGGCAGCCG ACCGCACAGA
	2151	TATGCCGGC ATCCGCCCCT ACGGCGCAAC TTTCCGCGCA GCGGCAGCCG
	2201	TACAGCATGC GAATGCCGCC GACGGTGTAC GCATCTTCAA CAGTCTCGCC
	2251	GCTACCGTCT ATGCCGACAG TACCGCCGC CATGCCGATA TGCAGGGACG
20	2301	CCGCCTGAAA GCGGTATCGG ACGGGTTGGG CCACAAACGGC ACGGGTCTGC
	2351	CGCTCATCGC GCAAAACCCAA CGAGCAGGGT GAACGTTGGG ACACGGGGT
	2401	GTTGAAGGCA AAATCGCCGG CAGTACCCAA ACCGTCGGCA TTGCGCGAA
	2451	AACCGGGCAA AATACGACAG CAGCCGCCAC ACTGGGCATG GGACGCAGCA
	2501	CATGGAGCGA AAACAGTGC AATGCAAAAA CCGACAGCAT TAGTCTGTT
	2551	GCAGGCATAC GGCACCGATGC GGGCGATATC GGCTATCTCA AAGGCTGTT
	2601	CTCCTACCGG CGCTACAAAAA ACAGCATCAG CGCGAGCAGC GGTCCGGACG
	2651	AACATCGGGA AGGCAGCGC AACGGCACGC TGATGCACTG GGGCGCAGTC
	2701	GGCGGTGTCG ACGTTCCGTT TGCGCGAACG GGAGATTG TGCTCGAAGG
	2751	CGGTCTGCGC TACCGACCTGC TCAAAACAGGA TGCAATTGCG GAAAAGGGCA
	2801	GTGCTTGGGG CTGGAGCGGC AACAGCCTCA CTGAAGGCAC GCTGGTCGGA
	2851	CTCGCGGGTC TGAAGCTGTC GCAACCCCTG AGCGATAAAAG CCGTCTGTT
	2901	TGCAACGGCG GCGCTGGAAC CGCACCTGAA CGGACCGCAG TACACGGTAA
	2951	CGGGCGGCTT TACCGCCGCG ACTGCGAGCA CGGGCAAGAC GGGGGCACGC
	3001	AATATGCCGC ACACCCGCT TGTTGCCGGG CTGGGCCGG ATGTCGAATT
	3051	CGGCAACGGC TGGAACCGGT TGGCACGTTA CAGCTACGCC GGTTCAAAC
	3101	AGTACGGCAA CCACAGCGGA CGAGTCGGCG TAGGCTACCG GTTCTCGAC
	3151	GGTGGCGGAG GCACTGGATC CTCAGATTG GCAAACGATT CTTTATCCG
30	3201	GCAGGTTCTC GACCGTCAGC ATTCGAACC CGACGGAAA TACCACCTAT
	3251	TCGGCAGCAG GGGGAACCTT GCGAGCGCAG CGGGCCATAT CGGATTGGGA
	3301	AAAATACAAA GCCATCAGT GGGCAACCTG ATGATTCAAC AGGGGGCCAT
	3351	TAAAGGAAT ATCGGTACAGA TTGTCCTGCT TTCCGATCAC GGGCACGAAG
	3401	TCCATTCCCC CTTCGACAAAC CATGCCCTCAC ATTCCGATT TGATGAAAGCC
	3451	GGTAGTCCCC TTGACGGATT TAGCCTTTAC CGCATCCATT GGACGGATA
40	3501	CGAACACCAC CCCGCCGACG GCTATGACGG GCCACAGGGC GGCGGTATC
	3551	CCGCTCCCAA AGGCGCGAGG GATATATACA GCTACGACAT AAAAGGCCTT
	3601	GCCCAAATAA TCCGCTCAA CTCGACCGAC AACCGCAGCA CCCGACAACG
	3651	GCTTGGCAGC CGTTCCACA ATGCCGGTAG TATGCTGACG CAAGGGAGTAG
	3701	GCGACGGATT CAAACGCGCC ACCCGATACA GCGCCGAGCT GGACAGATCG
	3751	GGCAATGCCG CCGAACGCTT CAACGGCACT GCAGATATCG TAAAAAACAT
	3801	CATCGGGCGC CGAGGAGAAA TTGTCGGCG AGGCGATGCC GTGCAGGGCA
	3851	TAAGCGAAGG CTAAACACATT GCTGTCATGC ACAGGTTGGG TCTGCTTCC
	3901	ACCGAAAACA AGATGGCGCG CATCAACGAT TTGGCAGATA TGCGCAACT
	3951	CAAAGACTAT GCGCCAGCAG CCATCCGCGA TTGGGCAGTC CAAAACCCCA
	4001	ATGCCGCACA AGGCATAGAA CGCGTCAGCA ATATCTTAT GGCAGCCATC
	4051	CCCATCAAAG GGATTGGAGC TGTCGGGGAA AAATACGGCT TGGCGGCAT
45	4101	CACGGCACAT CCTATCAAGC GGTGCGAGAT GGGCGCGATC GCATGCCGA
	4151	AAGGGAAATC CGCCGTCAGC GACAATTTCG CCGATGCCGC ATACGCCAA
	4201	TACCCGTCCT CTTACCATTC CGAATATTC CGTTCAAAC TGGAGCAGCG
	4251	TTACGGCAA AAAACATCA CCTCCTCAAC CGTGCCTGCCG TCAAACGGCA
	4301	AAAATGTCAA ACTGGCAGAC CAACGCCAAC CGAAGACAGG CGTACCGTTT
	4351	GACGGTAAAG GGTTCCGAA TTTGAGAAG CACGTAAAT ATGATACGCT
55	4401	CGAGCACCAC CACCACCAACTG

1 MTSAPDFNAG GTGIGSNSRA TTAKSAAVSY AGIKNEMCKD RSMLCAGRDD  
 51 VAVTDRDAKI NAPPNLHTG DFPNPNDAYK NLINLKPAIE AGYTGRGVEV  
 101 GIVDTGESVG SISFPELYGR KEHGYNENYK NYTAYMRKEA PEDGGGGK DIE  
 151 ASFDDEAVIE TEAKPTDIRH VKEIGHIDLV SHIIGGRSVD GRPAGGIAPD  
 201 ATLHIMNTND ETKNEMMVAA IRNAWVKLGE RGVRIVNNSF GTTSRAGTAD  
 251 LFQIANSEEQ YRQALLDYSG GDKTDEGIRL MQQSDYGNLS YHIRRNKNMLF  
 301 IFSTGNDAQA QPNTYALLPF YEKDAQKGII TVAGVDRSGE KFKREMYGEP  
 351 GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRTNPIQIAG TSFSAPIVTG  
 401 TAALLLQKYP WMSNDNLRTT LLTTAQDIGA VGVDSKFGWG LLDAGKAMNG  
 451 PASPFQGDF ADTKGTSDIA YSFRNDISGT GGLIKKGGSQ LQLHGNNTYT  
 501 GKTIIIEGGSL VLYGNNKSDM RVETKGALIY NGAASGGSLN SDGIVYLA DT  
 551 DQSGANETVH IKGSLQLDGK GTLYTRLGKL LKVDGTAIIIG GKLYMSARGK  
 601 GAGYLNSTGR RVPFLSAAKI GQDYSFFTNI ETDGGLLASL DSVEKTAGSE  
 651 GDTLSYYVRR GNAARTASAA AHSAPAGLKH AVEQGGSNLE NLMVELDASE  
 701 SSATPETVET AAADRTDMPG IRPYGATFRA AAAVQHANAA DGVrifnsla  
 751 ATVYADSTAA HADMQGRRLK AVSDGLDHNG TGLRVIAQTQ QDGGTWEQGG  
 801 VEGKMRGSTQ TVGIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF  
 851 AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHAEGSV NGTLMQLGAL  
 901 GGVNVPPFAAT GDLTVEGGLR YDLLKQDAFA EKGSALGWSG NSLTEGTLVG  
 951 LAGLKLSQPL SDKAVLFATA GVERDLNGRD YTWTGGFTGA TAATGKTGAR  
 1001 NMPHTRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLD  
 1051 GGGGTGSSDL ANDSFIRQVL DRQHFEPDGK YHLFGSRGEL AERSGHIGLG  
 1101 KIQSHQLGNL MIQQAAIKGN IGYIVRFSDH GHEVHSPFDN HASHSDSDEA  
 1151 GSPVTDGFSLY RIHWGDYEH PADGYDGPQG GGYPAPKGAR DIYSYDIKGV  
 1201 AQNIRLNLTD NRSTGQRLLAD RFHNAGSMLT QGVGDGFKRA TRYSPELDRS  
 1251 GNAAEAFTNGT ADIVKNIIGA AGEIVGAGDA VQGISEGSNI AVMHGLGLLS  
 1301 TENKMARIND LADMAQLKDY AAAAIRDWAV QNPNAAQGIE AVSNIFMAAI  
 1351 PIKGIGAVRG KYGLGGITAH PIKRSQMGA ALPKGKSAVS DNFADAAYAK  
 1401 YPSPYHSRNI RSNLEQRYGK ENITSSTVPP SNGKNVKLAD QRHPKTGVPF  
 1451 DGKGFPNFEK HVKYDTLEHH HHHH\*

30

35

40

45

50

55

AG983-741

5	1	ATGACTTCTG	CGCCCGACTT	CAATGCAGGC	GGTACCGGTA	TCGGCAGCAA
	51	CAGCAGAGCA	ACAACAGCGA	AATCAGCAGC	AGTATCTTAC	GCCGGTATCA
	101	AGAACGAAAT	GTGCAAAGAC	AGAAGCATGC	TCTGTGCCGG	TCGGGATGAC
	151	GTTGGGTTA	CAGACAGGGA	TGCCAAATC	AATGCCCCCC	CCCCGAATCT
	201	GCATACCGGA	GACTTCCAA	ACCCAAATGA	CGCATACAAG	AATTGATCA
	251	ACCTCAAACC	TGCAATTGAA	GCAGGCTATA	CAGGACGCGG	GGTAGAGGTA
	301	GGTATGCTCG	ACACAGGCGA	ATCCGTCGGC	AGCATATCCT	TTCCCGAACT
	351	GTATGGCAGA	AAAGAACACG	GCTATAACGA	AAATTACAAA	AACTATAACGG
10	401	CGTATATGCG	GAAGGAAGCG	CCTGAAGACG	GAGGCCGTAA	AGACATTGAA
	451	GCTTCTTCG	ACGATGAGGC	CGTTATAGAG	ACTGAAGCAA	AGCCGACGGA
	501	TATCCGCCAC	GTAAAAAGAAA	TCGGACACAT	CGATTGGTC	TCCCATATTA
	551	TTGGCGGGCG	TTCCGTGGAC	GGCAGACCTG	CAGGCCGTAT	TGCGCCCGAT
	601	GCGACGCTAC	ACATAATGAA	TACGAATGAT	GAAACCAAGA	ACGAAATGAT
15	651	GGTTGCAGCC	ATCCGCAATG	CATGGGTCAA	GCTGGCGAA	CGTGGCGTGC
	701	GCATCGTCAA	TAACAGTTTT	GGAACAAACAT	CGAGGGCAGG	CACTGCCGAC
	751	CTTTTCCAAA	TAGCCAATTC	GGAGGAGCAG	TACCGCCAAG	CGTTGCTCGA
	801	CTATTCCGGC	GGTGATAAAA	CAGACGAGGG	TATCCGCTG	ATGCAACAGA
	851	GCGATTACGG	CAACCTGTCC	TACCACATCC	GTAATAAAA	CATGCTTTTC
	901	ATCTTTTCGA	CAGGCAATGA	CGCACAAAGCT	CAGCCAACA	CATATGCCCT
20	951	ATTGCCATT	TATGAAAAAG	ACGCTAAAAA	AGGCATTATC	ACAGTCGCAG
	1001	GCGTAGACCG	CAGTGGAGAA	AAGTTCAAC	GGGAAATGTA	TGGAGAACCG
	1051	GGTACAGAAC	CGCTTGAGTA	TGGCTCAAAC	CATTGGGAA	TTACTGCCAT
	1101	GTTGGCCTG	TCGGCACCC	ATGAAGCAAG	CGTCCGTTTC	ACCCGTACAA
	1151	ACCCGATICA	AATTGCCGGA	ACATCCTTTT	CCGCACCCAT	CGTAACCGGC
	1201	ACGGCGGCTC	TGCTGCTGCA	GAAATACCCG	TGGATGAGCA	ACGACAACCT
25	1251	GCGTACCAACG	TTGCTGACGA	CGGCTCAGGA	CATCGGTGCA	GTCGGCGTGG
	1301	ACAGCAAGTT	CGGCTGGGGA	CTGCTGGATG	CGGGTAAGGC	CATGAACCGGA
	1351	CCCGCGTCCT	TTCCGTTCGG	CGACTTTACC	GCCGATACGA	AAGGTACATC
	1401	CGATATTGCC	TACTCCTTCC	GTAACGACAT	TTCAGGCACG	GGCGGCCTGA
	1451	TCAAAAAAGG	CGGCAGCCAA	CTGCAACTGC	ACGGCAACAA	CACCTATACG
	1501	GGCAAAACCA	TTATCGAAGG	CGGTTCGCTG	GTGTTGTACG	GCAACAAACAA
30	1551	ATCGGATATG	CGCGTCGAAA	CCAAAGGTGC	GCTGATTAT	AACGGGGCGG
	1601	CATCCGGCGG	CAGCCTGAAC	AGCGACGGCA	TTGTCTATCT	GGCAGATACC
	1651	GACCAATCCG	GCGCAAACGA	AACCGTACAC	ATCAAAGGCA	GTCTGCAGCT

35

40

45

50

55

	1701	GGACGGCAAA GGTACGCTGT ACACACGTTT GGGCAAACGT CTGAAAGTGG
5	1751	ACGGTACCGC GATTATCGGC GGCAAGCTGT ACATGTCGGC ACGCGGCAAG
	1801	GGGGCAGGCT ATCTAACAG TACCGGACGA CGTGTCCCT TCCTGAGTGC
	1851	CGCCAAAATC GGGCAGGATT ATTCTTCTT CACAAACATC GAAACCGACG
	1901	GCGGCCTGCT GGCTTCCCTC GACAGCGTCG AAAAAACAGC GGGCAGTGAA
	1951	GGCGACACGC TGTCCTATTA TGTCCGTGCG GGCAATGCGG CACGGACTGC
10	2001	TTCGGCAGCG GCACATTCCG CGCCCGCCGG TCTGAAACAC GCCGTAGAAC
	2051	AGGGCGGCAG CAATCTGGAA AACCTGATGG TCGAAGTGG A TGCCTCCGAA
	2101	TCATCCGCAA CACCCGAGAC GGTGAAACT GCGGCAGCCG ACCGCACAGA
	2151	TATGCCGGG ATCCGCCCT ACAGGGCAAC TTTCCCGCA GCGGCAGCCG
	2201	TACAGCATGC GAATGCCGC GACGGTGTAC GCATCTTCAA CAGTCTGCC
	2251	GCTACCGCT ATGCCGACAG TACCGCCGCC CATGCCGATA TGCAGGGACG
15	2301	CCGCCTGAAA GCCGTATCGG ACGGGTTGGA CCACAAACGGC ACGGGTCTGC
	2351	GCGTCATCGC GCAAACCCAA CAGGACGGTG GAACGTGGGA ACAGGGCGGT
	2401	GTTGAAGGCA AAATGCGCGG CAGTACCCAA ACCGTCGGCA TTGCCCGGAA
	2451	AACCGGCAA AATACGACAG CAGCCGCCAC ACTGGGCATG GGACGCAGCA
	2501	CATGGAGCGA AAACAGTGC AATGCAAAA CCGACAGCAT TAGTCTGTT
	2551	GCAGGCATAC GGCACGATGC GGGCGATATC GGCTATCTCA AAGGCCTGTT
20	2601	CTCCTACCGA CGCTACAAAA ACAGCATCAG CCGCAGCACC GTTGCAGGACG
	2651	AACATGCCGA AGGCAGCGTC AACGGCACGC TGATGCAGCT GGGCGCACTG
	2701	GGCGGTGTCA ACGTTCCGTT TGCCGCAACG GGAGATTGGA CGGTGAGAAG
	2751	CGGTCTGCGC TACCGACCTGC TCAAACAGGA TGCAATTGCC GAAAAGGCA
	2801	GTGCTTTGGG CTGGAGCGGC AACAGCCTCA CTGAAGGCAC GCTGGTCGGA
	2851	CTCGCGGTC TGAAGCTGTC GCAACCCCTG AGCGATAAAAG CCGTCTGTT
	2901	TGCAACGGCG GGGCGTGAAC GCGACCTGAA CGGACGCGAC TACACGGTAA
25	2951	CGGGCGGCTT TACCGGCGCG ACTGCAGCAA CCGGCAAGAC GGGGGCACGC
	3001	AATATGCCGC ACACCCGTCT GGTTGCCGGC CTGGGCGCGG ATGTCGAATT
	3051	CGGCAACGGC TGGAACGGCT TGGCACGTTA CAGCTACGCC GGTTCCAAAC
	3101	AGTACGGCAA CCACAGCGGA CGAGTCGGCG TAGGCTACCG GTTCTCGAG
	3151	GGATCCGGAG GGGGTGGTGT CGCCGCCGAC ATCGGTGCGG GGCTTGCCGA
30	3201	TGCACTAACCGC GCACCGCTCG ACCATAAAAGA CAAAGGTTG CAGTCTTGA
	3251	CGCTGGATCA GTCCGTCAAGG AAAAACGAGA AACTGAGCT GCGGGCACAA
	3301	GGTGCAGAAA AAACCTTATGG AAACGGTGAC AGCCTCAATA CGGGCAAATT
	3351	GAAGAACGAC AAGGTAGGCC GTTTCGACT TATCCGCAA ATCGAAGTGG
	3401	ACGGGCAGCT CATTACCTTG GAGAGTGGAG AGTCCAAGT ATACAAACAA
	3451	AGCCATTCCG CCTTAACCGC CTTTCAGACC GAGCAAATAC AAGATTGCGA
35	3501	GCATTCCGGG AAGATGGTTG CGAAACGCCA GTTCAGAACATC GCGCACATAG
	3551	CGGGCGAACAA TACATCTTT GACAAGCTTC CGAAGGGCGG CAGGGCGACA
	3601	TATCGGGGA CGGCCTTCGG TTCAGACGAT GCCGGGGAA AACTGACCTA
	3651	CACCATAGAT TTCCGCCAAC AGCAGGGAAA CGGCAAATC GAACATTGA
	3701	AATCGCCAGA ACTCAATGTC GACCTGGCG CCGCCGATAT CAAGCCGGAT
	3751	GGAAAACGCC ATGCCGTAT CAGCGGTTCC GTCTTTACA ACCAAGCCGA
40	3801	GAAAGGCACT TACTCCCTCG GTATCTTGG CGGAAAAGCC CAGGAAGTTG
	3851	CCGGCAGCGC GGAAGTGAAA ACCGTAAACG GCATACGCCA TATCGGCCTT
	3901	GCGGCCAACG AACTCGAGCA CCACCAACAC CACCAACTGA

45

50

55

1 MTSAPDFNAG GTGIGSNSRA TTAKSAAVSY AGIKNEMCKD RSMLCAGRDD  
 51 VAVTDRDAKI NAPPNLHTG DFPNPNDAYK NLINLKPAIE AGYTGRGV  
 101 GIVDTGESVG SISFPELYGR KEHGNENYK NYTAYMRKEA PEDGGKDIE  
 151 ASFDDEAVIE TEAKPTDIRH VKEIGHIDLV SHIIGGRSVD GRPAGGIAPD  
 201 ATLHIMNTND ETKNEMMVAI IRNAWVKLGE RGVRIVNNSF GTTSRAGTAD  
 251 LFQIANSEEQ YRQALLDYSG GDKTDEGIRL MQQSDYGNLS YHIRNKNMLF  
 301 IFSTGNDQAQ QPNTYALLPF YEKDAQKGII TVAGVDRSGE KFKREMYGEP  
 351 GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRTNPIQIAG TSFSAPIVTG  
 401 TAALLLQKYP WMSNDNLRTT LLTTAQDIAGA VGVDSKFGWG LLDAGKAMNG  
 451 PASFPFGDFT ADTKGTSIA YSFRNDISGT GGLIKGGSQ LQLHGNNTYT  
 501 GKTIIEGGSL VLYGNNKSDM RVETKGALIY NGAASGGSLN SDGIVYLADT  
 551 DQSGANETVH IKGSLQLDGK GTLYTRLGKL LKVDGTAIIIG GKLYMSARGK  
 601 GAGYLNSTGR RVPFLSAAKI QGDYSFFTNI ETDGGLLASL DSVEKTAGSE  
 651 GDTLSYYVRR GNAARTASAA AHSAPAGLKH AVEQGGSNLE NLMVELDASE  
 701 SSATPETVET AAADRTDMPG IRPYGATFRA AAAVQHANAA DGVRIFNSLA  
 751 ATVYADSTA HADMQGRRLK AVSDGLDHNG TGLRVIAQTQ QDGTTWEQGG  
 801 VEGKMRGSTQ TVGIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF  
 851 AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHAEGSV NGTLMQLGAL  
 901 GGVNVPFAT GDLTVEGGLR YDLLKQDFAA EKGSLGWG NSLTEGTLVG  
 951 LAGLKLSQPL SDKAVLFATA GVERDLNGRD YTWTGGFTGA TAATGKTGAR  
 1001 NMPTHRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLE

20

1051 GSGGGGVAAAD IGAGLADALT APLDHDKGL QSLTLDQSVR KNEKLKLAAQ  
 1101 GAEKTYGNGD SLNTGKLND KVSRFDFIRQ IEVDGQLITL ESGEFQVYKQ  
 1151 SHSALTAFQT EQIQDSEHSG KMVAKRQFRI GDIAGEHTSF DKLPEGGRAT  
 1201 YRGTAFGSDD AGGKLTYTID FAAKQGNGKI EHLKSPELNV DLAAADIKPD  
 1251 GKRHAVISGS VLYNQAEKGS YSLGIFGGKA QEVAGSAEVK TVNGIRHIGL  
 1301 AAKQLEHHHH HH\*

30

35

40

45

50

55

**ΔG983-961**

5	1	ATGACTTCTG	CGCCCGACTT	CAATGCAGGC	GGTACCGGTA	TGGCAGCAA
	51	CAGCAGAGCA	ACAACAGCGA	AATCAGCAGC	AGTATCTTAC	GCGGTATCA
	101	AGAACGAAAT	GTGCAAAGAC	AGAACGATGC	TCTGTGCCGG	TCGGGATGAC
	151	GTTCCGGTTA	CAGACAGGGGA	TGCCAAAATC	AATGCCCCC	CCCCGAATCT
	201	GCATACCGGA	GACTTTCAA	ACCCAAATGA	CGCATACAAG	AATTGATCA
	251	ACCTCAAACC	TGCAATTGAA	GCAGGCTATA	CAGGACGCCG	GGTAGAGGTA
	301	GGTATCGTCG	ACACAGGCAG	ATCCGTCGGC	AGCATATCCT	TTCCCGAACT
	351	GTATGGCAGA	AAAGAACACG	GCTATAACGA	AAATTACAAA	AACTATACGG
10	401	CGTATATGCG	GAAGGAAGCG	CCTGAAGACG	GAGGGCGTAA	AGACATTGAA
	451	GCTTCTTCG	ACGATGAGGC	CGTTATAGAG	ACTGAAGCAA	AGCCGACGGA
	501	TATCCGCCAC	GTAAAAGAAA	TCGGACACAT	CGATTGGTC	TCCCATATTA
	551	TTGGCAGGGCG	TTCCGTGGAC	GGCAGACCTG	CAGGGCGTAT	TGCGCCCGAT
	601	GCGACGCTAC	ACATAATGAA	TACGAATGAT	GAAACCAAGA	ACGAAATGAT
15	651	GGTTGCAGCC	ATCCGCAATG	CATGGGTCAA	GCTGGGCGAA	CGTGGCGTGC
	701	GCATCGTCAA	TAACAGTTTT	GGAAACAACAT	CGAGGGCAGG	CACTGCCGAC
	751	CTTTTCCAAA	TAGCCAATTC	GGAGGAGCAG	TACCGCCAAG	CGTTGCTCGA
	801	CTATTCCGGC	GGTGATAAAA	CAGACGAGGG	TATCCGCCTG	ATGCAACAGA
	851	GCGATTACGG	CAACCTGTCC	TACCACATCC	GTAATAAAAA	CATGCTTTTC
	901	ATCTTTCGA	CAGGCAATGA	CGCACAAAGCT	CAGCCAACA	CATATGCCCT
20	951	ATTGCCATT	TATGAAAAAG	ACGCTAAAAA	AGGCATTATC	ACAGTCGCA
	1001	GCGTAGACCG	CAGTGGAGAA	AAGTTCAAAC	GGGAAATGTA	TGGAGAACCG
	1051	GGTACAGAAC	CGCTTGAGTA	TGGCTCCAAC	CATTGCGGAA	TTACTGCCAT
	1101	GTGGTGCCTG	TCGGCACCCCT	ATGAAGCAAG	CGTCGCTTTC	ACCCGTACAA
	1151	ACCCGATTCA	AATTGCCGGA	ACATCCTTT	CCGCACCCAT	CGTAACCGGC
	1201	ACGGCGGCTC	TGCTGCTGCA	GAAATACCCG	TGGATGAGCA	ACGACAACCT
25	1251	GGCTTACACG	TTGCTGACGA	CGGCTCAGGA	CATCGGTGCA	GTCGGCGTGG
	1301	ACAGCAAGTT	CGGGTGGGGA	CTGCTGGATG	CGGGTAAGGC	CATGAACGGG
	1351	CCCGCGTCC	TTCCGTTCCG	CGACTTTACC	GGCGATACGA	AAGGTACATC
	1401	CGATATTGCC	TACTCCCTCC	GTAACGACAT	TTCAGGCACG	GGCAGGCTGA
	1451	TCAAAAAAAGG	CGGCAGCCAA	CTGCAACTGC	ACGGCAACAA	CACCTATACG
30	1501	GGCAAAACCA	TTATCGAAGG	CGGTTCGCTG	GTGTTGTACG	GCAACAACAA
	1551	ATCGGATATG	CGCGTCGAAA	CCAAAGGTGC	GCTGATTAT	AACGGGGCGG
	1601	CATCCGGCGG	CAGCCTGAAC	AGCGACGGCA	TTGTCTATCT	GGCAGATACC
	1651	GACCAATCCG	GCGCAAACGA	AACCGTACAC	ATCAAAGGCA	GTCTGCAGCT
	1701	GGACGGGAAA	GGTACGCTGT	ACACACGTT	GGGCAAACTG	CTGAAAGTG
	1751	ACGGTACGGC	GATTATCGGC	GGCAAGCTGT	ACATGTCGGC	ACGCGGCAAG
35	1801	GGGGCAGGCT	ATCTCAACAG	TACCGGACGA	CGTGTCCCT	TCCTGAGTGC
	1851	CGCCAAAATC	GGGCAGGATT	ATTCTTTCTT	CACAAACATC	GAAACCGACG
	1901	GGGGCCTGCT	GGCTTCCCTC	GACAGCGTCG	AAAAAACAGC	GGGCAGTGAA
	1951	GGGCACACGC	TGTCCTATT	TGTCCGTCGC	GGCAATGCGG	CACGGACTGC
	2001	TTCGGCAGCG	GCACATTCCG	CGCCCGCCGG	TCTGAAACAC	GCCGTAGAAC
	2051	AGGGCGGCAG	CAATCTGGAA	AACCTGATGG	TCGAACCTGGA	TGCCTCCGAA
40	2101	TCATCCGCAA	CACCGAGAC	GGTTGAAACT	GGGGCAGCCG	ACCGCACAGA
	2151	TATGCCGGGC	ATCCGCCCCCT	ACGGCGCAAC	TTTCCGCGCA	GCGGCAGCCG
	2201	TACAGCATGC	GAATGCCGCC	GACGGTGTAC	GCATCTTCAA	CAGTCTCGCC
	2251	GCTACCGTCT	ATGCCGACAG	TACCGCCGCC	CATGCCGATA	TGCAGGGACG
	2301	CCGCTGAAA	GGCGTATCGG	ACGGGTTGGA	CCACAAACGGC	ACGGGTCTGC
	2351	GCGTCATCGC	GCAAACCCAA	CAGGACGGTG	GAACGTGGGA	ACAGGGCGGT
45	2401	GTTGAAGGCA	AAAATGCGCG	CAGTACCCAA	ACCGTCGGCA	TTGCCCGCAA
	2451	AACCGGCAGA	AATAACGACAG	CAGCCGCCAC	ACTGGGCATG	GGACGCAGCA
	2501	CATGGAGCGA	AAACAGTGC	AATGCAAAAA	CCGACAGCAT	TAGTCTGTTT
	2551	GCAGGCATAC	GGCACGATGC	GGGCATATC	GGCTATCTCA	AAGGCCTGTT
	2601	CTCCTACGGA	CGCTACAAAA	ACAGCATCAG	CCGCAGCACC	GGTGCAGGACG
	2651	AACATGCCGA	AGGCAGCGTC	AACGGCAGCG	TGATGCACTG	GGGCGCACTG
	2701	GGCGGTGTCA	ACGTTCCGTT	TGCGCAACG	GGAGATTG	CGGTGCAAGG
	2751	CGGTCTGCC	TACGACCTGC	TCAAACAGGA	TGCATTCGCC	GAAAAGGCA
	2801	GTGCTTTGGG	CTGGAGCGGC	AACAGCCTCA	CTGAAGGCAC	GCTGGTCGGA
50	2851	CTCGCGGGTC	TGAAGCTGTC	GCAACCCCTG	AGCGATAAAAG	CCGTCCCTGTT

**EP 1 790 660 A2**

	2901	TGCAACGGCG	GGCGTGAAC	GCGACCTGAA	CGGACGCGAC	TACACGGTAA
5	2951	CGGGCGGCTT	TACCGCGCG	ACTGCAGCAA	CCGGCAAGAC	GGGGCACGC
	3001	AATATGCCGC	ACACCCGTCT	GGTTGCCGGC	CTGGGCGCG	ATGTCGAATT
	3051	CGGCAACGGC	TGGAACGGCT	TGGCACGTTA	CAGCTACGCC	GGTTCCAAAC
	3101	AGTACGGCAA	CCACAGCGGA	CGAGTCGGCG	TAGGCTACCG	GTTCCCTCGAG
	3151	GGTGGCGGAG	GCACGGATC	CGCCACAAAC	GACGACGATG	TTAAAAAAAGC
10	3201	TGCCACTGTC	GCCATTGCTG	CTGCCTACAA	CAATGGCCA	AAATCAACG
	3251	GTTTCAAAGC	TGGAGAGACC	ATCTACGACA	TTGATGAAGA	CGGCACAAATT
	3301	ACCAAAAAG	ACGCAACTGC	AGCCGATGTT	GAAGCCGACG	ACTTTAAAGG
	3351	TCTGGGTCTG	AAAAAAAGTCG	TGACTAACCT	GACCAAAACC	CTCAATGAAA
	3401	ACAAACAAAA	CGTCGATGCC	AAAGTAAAAG	CTGCAGAACATC	TGAAATAGAA
	3451	AAAGTTAACAA	CCAAGTTAGC	AGACACTGAT	GCCGCTTCTAG	CAGATACTGA
	3501	TGCCGCTCTG	GATGCAACCA	CCAACGCCCTT	GAATAAAATTG	GGAGAAAATA
	3551	TAACGACATT	TGCTGAAGAG	ACTAAGACAA	ATATCGTAAA	AATTGATGAA
15	3601	AAATTAGAAG	CCGTGGCTGA	TACCGTCGAC	AAGCATGCCG	AAGCATTCAA
	3651	CGATATCGCC	GATTCAATTGG	ATGAAACCAA	CACTAAGGCA	GACGAAGCCG
	3701	TCAAAACCGC	CAATGAAGCC	AAACAGACGG	CCGAAGAAC	CAACACAAAC
	3751	GTCGATGCCA	AAAGTAAAAGC	TGCAGAAACT	GCAGCAGGCA	AAGCCGAAGC
	3801	TGCCGCTGGC	ACAGCTAATA	CTGCAGCGGA	CAAGGCCGAA	GCTGTCGCTG
	3851	CAAAGGTTAC	CGACATCAA	GCTGATATCG	CTACGAACAA	AGATAATATT
20	3901	GCTAAAAAAAG	CAAACAGTGC	CGACGTGTAC	ACCAGAGAAG	AGTCTGACAG
	3951	CAAATTGTC	AGAATTGATG	GTCTGAACGC	TACTACCGAA	AAATTGGACA
	4001	CACGCTTGGC	TTCTGCTGAA	AAATCCATTG	CCGATCACGA	TACTCGCCTG
	4051	AACGGTTTGG	ATAAAACAGT	GTCAGACCTG	CGCAAAGAAA	CCCGCCAAGG
	4101	CCTTGCAGAA	CAAGCCGCGC	TCTCCGGTCT	GTTCCAACCT	TACAACGTGG
25	4151	GTCGGTTCAA	TGTAACGGCT	GCAGTCGGCG	GCTACAAATC	CGAACATGGCA
	4201	GTCGCCATCG	GTACCGGCTT	CCGCTTTACC	AAAAACTTG	CCGCCAAAGC
	4251	AGGC GTGGCA	GTGGCACTT	CGTCCGGTTC	TTCCGAGCC	TACCATGTCG
	4301	CGCTCAATTA	CGAGTGGCTC	GAGCACCACC	ACCACCA	CTGA
30	1	MTSAPDFNAG	GTGIGSNSRA	TTAKSAAVSY	AGIKNEMCKD	RSMILCAGRDD
	51	VAVTDRDAKI	NAPPNLHTG	DFPNPNDAYK	NLINLKPAIE	AGYTGRGVEV
	101	GIVDTGESVG	SISFPELYGR	KEHGYNENYK	NYTAYMRKEA	PEDGGGKDIE
	151	ASFDEAVIE	TEAKPTDIRH	VKEIGHIDLV	SHIIGGRSVD	GRPAGGIAPD
	201	ATLHIMNTND	ETKNEMMVAI	IRNAWVKLGE	RGVRIVNNSF	GTTSRAGTAD
35	251	LFQIANSEEQ	YRQALLDYSG	GDKTDEGIRL	MQQSDYGNLS	YHIRKNMLF
	301	IFSTGNDQAQ	QPNTYALLPF	YEKDAQKGII	TVAGVDRSRE	KFKREMYGEP
	351	GTEPLEYGSN	HCGITAMWCL	SAPYEASVRF	TRTNPIQIAG	TSFSAPIVTG
	401	TAALLLQKYP	WMSNDNLRTT	LLTTAQDIGA	VGVDSDKFGWG	LLDAKGAMNG
	451	PASFPFGDF	ADTKGTSdia	YSFRNDISGT	GGLIKGGSQ	LQLHGNNTYT
	501	GKTIEEGSSL	VLYGNNKSDM	RVETKGALIY	NGAASGGSLN	SDGIVYLADT
40	551	DQSGANETHVH	IKGSLQLDGK	GTLYTRLGKL	LKVDTAIIG	GKLYMSARGK
	601	GAGYLNSTGR	RVPFLSAAKI	QDODYSFFTNI	ETDGGLLASL	DSVEKTAGSE
	651	GDTLSYYVRR	GNAARTASAA	AHSAPAGLKH	AVEQGGSNLE	NLMVELDASE
	701	SSATPETVET	AAADRTDMPG	IRPYGATFRA	AAA VQHANAA	DGVRIFNLA
	751	ATVYADSTAA	HADMQGRRLK	AVSDGLDHNG	TGLRVIAQTQ	QDGGTWEQGG
45	801	VEGMGRGSTQ	TVGIAAKTGE	NTTAAATLGM	GRSTWSENSA	NAKTDISL
	851	AGIRHDAGDI	GYLKGLFSYG	RYKNSISRST	GADEHAEGSV	NGTLMQJGAL
	901	GGVNVPFAAT	GDLTVEGGGLR	YDLLKQDFA	EKGSLGWG	NSLTEGTLVG
	951	LAGLKLSQLP	SDKAVLFATA	GVERDLNGRD	YTVTGGFTGA	TAATGKTGAR
	1001	NMPHTRLVAG	LGADVEFGNG	WNGLARYSYA	GSKQYGNHSG	RVGVGYRFLE
	1051	GGGGTGSATN	DDDVKKAATV	AIAAAYNNQ	EINGFKAGET	IYDIDEDGTI
50	1101	TKKDATAADV	EADDFKGLGL	KKVVTNLTKT	VNEKQNVDA	KVKAEESEIE
	1151	KLTTKLADTD	AALADTDAAL	DATTNALNKL	GENITTFAEE	TKTNIVKIDE
	1201	KLEAVADTV	KHAEAFNDIA	DSLDETNKA	DEAVKTANE	KOTAEETKQN
	1251	VDAKVKAAT	AAGKAEAAAG	TANTAADKAE	AVAAKVTDIK	ADIA TNKDNI
	1301	AKKANSADVY	TREESDSKFV	RIDGLNATTE	KLDTRLASAE	KSIADHDTRL
	1351	NGLDKTVSDL	RKETRQGLAE	QAALSGLFQP	YNVGRFNVTA	AVGGYKSES
	1401	VAIGTGFRFT	ENFAAKAGVA	VGTSSGSSAA	YHVGVNYEWL	EEHHHHHH*

**ΔG983-961c**

5            1 ATGACTTCTG CGCCCGACTT CAATGCAGGC GGTACCGGTA TCGGCAGCAA  
51          CAGCAGAGCA ACAACAGCGA AATCAGCAGC AGTATCTTAC GCCGGTATCA  
101         AGAACGAAAT GTGCAAAGAC AGAACGATGC TCTGTGCCGG TCGGGATGAC  
151         GTTGCAGGTTA CAGACAGGGGA TGCCAAAATC AATGCCCCCCC CCCCGAATCT  
201         GCATACCGGA GACTTTCCAA ACCCAAATGA CGCATACAAG AATTGATCA

10

15

20

25

30

35

40

45

50

55

	251	ACCTCAAACC	TGCAATTGAA	GCAGGGCTATA	CAGGACGCCG	GGTAGAGGTA
5	301	GGTATCGTCG	ACACAGGCAGA	ATCCGTGCGC	AGCATATCCT	TTCCCGAAGT
	351	CTATGGCAGA	AAAAGAACACG	GCTATAACGA	AAATTACAAA	AACTATAACGG
	401	CGTATATGCG	GAAGGAGAGC	CCTGAAGACG	GAGGCGGTAA	AGACATTGAA
	451	GCTTCTTCG	ACGATGAGGC	CGTTATAGAG	ACTGAAGCAA	AGCCGACGGA
	501	TATCCGCCAC	GTAAAAAGAAA	TCGGACACAT	CGATTGGTC	TCCCATATTA
	551	TTGGCGGGCG	TTCCGTGGAC	GGCAGACCTG	CAGGCGGTAT	TGCGCCCGAT
	601	GCGACGCTAC	ACATAATGAA	TACGAATGAT	GAAACCAAGA	ACGAAATGAT
	651	GGTTGAGCAG	ATCCGCAATG	CATGGGTCAA	GCTGGCGAA	CGTGGCGTGC
	701	GCATCGCTAA	TAACAGTTT	GGAAACAACAT	CGAGGGCAGG	CACTGCCCAC
10	751	CTTTTCCAAA	TAGCCTAATT	GGAGGAGCAG	TACGCCAAG	CGTGTCTCGA
	801	CTATTCGGC	GGTGATAAAA	CAGACGAGGG	TATCCGCTG	ATGCAACAGA
	851	GCGATTACGG	CAACCTGTCC	TACCACATCC	GTAATAAAA	CATGCTTTTC
	901	ATCTTTTCGA	CAGGAATGA	CGCACAAAGCT	CAGCCAAACA	CATATGCCCT
	951	ATTGCCATT	TATGAAAAAG	ACGCTAAAAA	AGGCATTATC	ACAGTCGAG
15	1001	GCGTAGACCG	CACTGGAGAA	AAGTTCAAAC	GGGAATGTA	TGGAGAACCG
	1051	GGTACAGAAC	CGCTTGAGTA	TGGCTCAAAC	CATTGCGGAA	TTACTGCCAT
	1101	GTGGTCTG	TCGGCACCC	ATGAAGCA	CGTCCGTTTC	ACCCGTACAA
	1151	ACCCGATTCA	AATTGCCGGA	ACATCCTTT	CCGACCCAT	CGTAACCGGC
	1201	ACGGCGGCTC	TGCTGCTGCA	GAAATACCCG	TGGATGAGCA	ACGACAACCT
	1251	GCGTACCAACG	TTGCTGACGA	CGGCTCAGGA	CATCGGTGCA	GTCGGCGTGG
	1301	ACAGCAAGTT	CGGCTGGGA	CTGCTGGATG	CGGGTAAGGC	CATGAACGGA
20	1351	CCCGCGTCT	TTCCGTTCGG	CGACTTTACC	GCCGATACGA	AAGGTACATC
	1401	CGATATTGCC	TACTCCTTCC	GTAACGACAT	TTCAGGCACG	GGCCGCTGAA
	1451	TCAAAAAAAGG	CGCGAGCCAA	CTGCAACTGC	ACGGCAACAA	CACCTATACG
	1501	GGCAAAACCA	TTATCGAAGG	CGGTTCGCTG	GTGTTGTAACG	GCAACAACAA
	1551	ATCGGATATG	CGCGTCGAAA	CCAAAGGTGC	GCTGATTAT	AACGGGGCGG
	1601	CATCCGGCG	CAGCCTGAAC	AGCGACGGCA	TTGTCTATCT	GGCAGATAAC
25	1651	GACCAATCCG	GGCCAAACGA	AACCGTACAC	ATCAAAGGA	GTCCTGCAGCT
	1701	GGACGCGAAA	GGTACGCTGT	ACACACGTT	GGGCAAACTG	CTGAAAGTGG
	1751	ACGGTACGGC	GATTATCGGC	GGCAAGCTGT	ACATGTCGGC	ACCGGGCAAG
	1801	GGGGCAGGCT	ATCTAACAG	TACCGGACGA	CGTGTCCCT	TCCTGAGTGC
	1851	CGCCAAAATC	GGGCAGGATT	ATTCTTTCTT	CACAAACATC	GAAACCGACG
	1901	GGGGCCTGCT	GGCTTCCCTC	GACAGCGTCG	AAAAAACAGC	GGGCAGTGA
30	1951	GGCGACACGC	TGTCTATT	TGTCCGTCGC	GGCAATGCGG	CACGGACTGC
	2001	TTCGGCACCG	GCACATTCCG	CGCCCCGCCG	TCTGAAACAC	GCCGTAGAAC
	2051	AGGGCGGC	CAATCTGGAA	AACCTGATGG	TCGAACCTGGA	TGCTCCGAA
	2101	TCATCCGCAA	CACCCGGAGC	GGTTGAAACT	CGGGCAGCCG	ACCGCACAGA
	2151	TATGCCGGC	ATCCGCCCC	ACGGCGCAAC	TTTCCGCGCA	GGGGCAGCCG
	2201	TACAGCATGC	GAATGCCGCC	GACGGTGTAC	GCATCTCAA	CAGTCTCGCC
	2251	GCTACCGTCT	ATGCCGACAG	TACCGCCGCC	CATGCCGATA	TGCAAGGACG
35	2301	CCGCCTGAAA	GGCGTATCGG	ACGGGTTGGA	CCACAACGGC	ACGGGTCTGC
	2351	GCGTCATCGC	GCAAAACCCAA	CAGGACGGTG	GAACGTGGG	ACAGGGCGGT
	2401	GTTGAAGCCA	AAATGCGCGG	CAGTACCCAA	ACCGTCGGCA	TTGCCGCGAA
	2451	AACCGGGAA	AATACGACAG	CAGCCGCCAC	ACTGGGCATG	GGACGCAGCA
	2501	CATGGAGCGA	AAACAGTGC	AATGCAAAA	CCGACAGCAT	TAGTCTGTT
	2551	GCACCCATAC	GGCACGATGC	GGGGCATAATC	GGCTATCTCA	AAGGCCTGTT
40	2601	CTCCTACGGA	CGCTACAAAA	ACAGCATCAG	CCGCAGCACC	GGTGCAGACG
	2651	AACATCGGAA	AGGCAGCGTC	AACGGCACGC	TGATGCAGCT	GGGCGCACTG
	2701	GGCGGGTCTA	ACGTTCCGTT	TGCGCAACG	GGAGATTGTA	CGGTGAGAAGG
	2751	CGGTCTGCG	TACGACATG	TCAAACAGGA	TGCGATTCGCC	GAAAAAGGCA
	2801	TGTGTTGGG	CTGGAGCGGC	AACAGCCTCA	TGAAGGGCAC	GCTGGTCGGA
	2851	CTCGGGGTC	TGAAGCTGTC	GCAACCTTCA	AGCGATAAAG	CCGTCCTGTT
	2901	TGCAACGGCG	GGCGTGGAA	GCGACCTGAA	CGGACCGCAG	TACACGGTAA
	2951	CGGGCGGCTT	TACCGGGCG	ACTGCAGCAA	CCGGCAAGAC	GGGGGCACGC
	3001	AATATGCCG	ACACCCGTCT	GGTTGCCGCG	CTGGGCGCG	ATGTCGAATT
	3051	CGGCAACGGC	TGGAACCGGC	TGGCACGTTA	CAGCTACGCC	GGTCCAAAC
	3101	ACTACCCGAA	CCACACGGGA	CGAGTCGGCG	TAGGCTACCC	GTTCCTCGAG
	3151	GGTGGCGGAG	GCATGGGATC	CGCCACAAAC	GACGACGATG	TTAAAAAAGC
45	3201	TGCCACTGTG	GCCATTGCTG	CTGCCCTACAA	CAATGGCAA	GAATCAACG
	3251	GTTTCAAAGC	TGGAGAGACC	ATCTACGACA	TTGATGAAAGA	CGGCACAATT
	3301	ACCAAAAAAG	ACGCAACTGC	AGCCGATGTT	GAAGCCGACG	ACTTTAAAGG
	3351	TCTGGGTCTG	AAAAAAAGTCG	TGACTAACCT	GACCAAAACC	GTCAATGAAA
	3401	ACAAACAAAA	CGTCGATGCC	AAAGTAAAAG	CTGCAAGAATC	TGAAATAGAA
	3451	AAGTTAACAA	CCAAGTTAGC	AGACACTGAT	GCCGCTTGTAG	CAGATACTGA
50	3501	TGCCGCTCTG	GATGCAACCA	CCAACGCCCT	GAATAAATTG	GGAGAAAATA
	3551	TAACGACATT	TGCTGAAGAG	ACTAAGACAA	ATATCGTAA	AATTGATGAA

3601 AAATTAGAAG CCGTGGCTGA TACCGTCGAC AAGCATGCCG AAGCATTCAA  
 3651 CGATATCGCC GATTCAATTGG ATGAAACCAA CACTAAGGCA GACGAAGCCG  
 3701 TCAAAACCAC CAATGAAGCC AAACAGACGG CCGAAGAAC CAAACAAAAC  
 3751 GTCGATGCCA AAGTAAAAGC TGCAAGAAACT GCAGCAGGCA AAGCCGAAGC  
 5 3801 TGCGCTGGC ACAGCTAATA CTGCAGCCG AAGGCCGAA GCTGTCGCTG  
 3851 CAAAAGTTAC CGACATCAA GCTGATATCG CTACCGAAC AGATAATATT  
 3901 GCTAAAAAAG CAAACAGTGC CGACGTGTAC ACCAGAGAAC AGTCTGACAG  
 3951 CAAATTGTC AGAATTGATG GTCTGAACGC TACTACCGAA AAATTGGACA  
 4001 CACGCTTGGC TTCTGCTGAA AAATCCATTG CCGATCACGA TACTCGCCTG  
 10 4051 AACGGTTTGG ATAAAACAGT GTCAGACCTG CGCAAAGAAA CCCGCCAAGG  
 4101 CCTTGAGAA CAAGCCGCGC TCTCCGGTCT GTTCCAACCT TACAACGTGG  
 4151 GTCTCGAGCA CCACCACCA CACCACTGA

15 1 MTSAPDFNAG GTGIGSNSRA TTAKSAAVSY AGIKNEMCKD RSMLCAGRDD  
 51 VAVTDRDAKI NAPPNLHTG DFPNPNDAYK NLNIKPAIE AGYTGRGVVEV  
 101 GIVDTGESVG SISFPELYGR KEHGYNENYK NYTAYMRKEA PEDGGGKDIE  
 151 ASFDEAVIE TEAKPTDIRH VKEIGHIDLV SHIIGGRSVD GRPAGGIAPD  
 201 ATLHIMNTND ETKNEMMVA IRNAWVKLGE RGVRIVNNSF GTTSRAGTAD  
 251 LFQIANSEQ YRQALLDYSG GDKTDEGIRL MQQSDYGNL YHIRRNKNMLF  
 301 IFSTGNDQA QPNTYALLPF YEKDAQKGII TVAGVDRSGE KFKREMYGEP  
 351 GTEPLEYGSN HCGITAMWCL SAPYEASVRF TRTNPIQIAQ TSFSAPIVTG  
 401 TALLLQKYP WMSNDNLRTT LLTTAQDIGA VGVDSKFCWG LLDAGKAMNG  
 451 PASFPFGDFT ADTKGTSdia YSFRNDISGT GGLIKKGGSQ LQLHGNNTYT  
 501 GKTIIEGGSL VLYGNNSDM RVETKGALIY NGAASGGSLN SDGIVYLADT  
 551 DQSGANETVH IKGSLQLDGK GTLYTRLGKL LKVDTAIIG GKLYMSARGK  
 601 GAGYLNSTGR RVPFLSAKI GQDYSFFTNI ETDGLLASL DSVEKTAGSE  
 651 GDTLSYYVRR GNAARTASAA AHSAPAGLKH AVEQGGSNLE NLMVELDASE  
 701 SSATPETVET AAADRTDMPG IRPYGATFRA AAAVQHANAA DGVRIFNSLA  
 751 ATVYADSTAA HADMQGRRLK AVSDGLDHNG TGLRVIAQTQ QDGGTWEQGG  
 801 VEGKMRGSTQ TVGIAAKTGE NTTAAATLGM GRSTWSENSA NAKTDSISLF  
 851 AGIRHDAGDI GYLKGLFSYG RYKNSISRST GADEHAEGSV NGTLMQLGAL  
 901 GGVNVPFAAT GDLTVEGGLR YDLLKQDRAFTA EKGSLGWSG NSLTEGTLVG  
 951 LAGLKLSQLP SDKAVLFATA GVERDLNGRD YTWTGGFTGA TAATGKTGAR  
 1001 NMPHTRLVAG LGADVEFGNG WNGLARYSYA GSKQYGNHSG RVGVGYRFLE  
 1051 GGGGTGSATN DDDVKKAATV AIAAAYNNQQ EINGFKAGET IYDIDEDGTI  
 1101 TKKDATAADV EADDFKGLGL KKVVNTLKT VNENKQNVD KVKAEESEIE  
 35 1151 KLTTKLAQTD AALADTDAAL DATTNALNKL GENITFAEE TKTNIVKIDE  
 1201 KLEAVADTVK KHAEAFTNDIA DSLDETNTKA DEAVKTANEK KOTAEETKON  
 1251 VDAKVKAET AAGKAEAAAG TANTAADKAE AVAAKVTDIK ADIATNKDNI  
 1301 AKKANSADV TREESDSKFV RIDGLNATTE KLDTRLASAE KSIADHDTRL  
 1351 NGGLDKTVSDL RKETRQGLAE QAALSGLFQP YNVGLEHHHH HH\*

40

*ΔG741 and hybrids*

**[0225]** Bactericidal titres generated in response to ΔG741 (His-fusion) were measured against various strains, including the homologous 2996 strain:

45

	2996	MC58	NGH38	F6124	BZ133
ΔG741	512	131072	>2048	16384	>2048

50

**[0226]** As can be seen, the ΔG741-induced anti-bactericidal titre is particularly high against heterologous strain MC58.

**[0227]** ΔG741 was also fused directly in-frame upstream of proteins 961, 961c, 983 and ORF46.1:

55

AG741-961

5                   1 ATGGTCGCGC CCGACATCGG TGCGGGGCCTT GCCGATGCAC TAACCGCAC  
      51 GCTCGACCAT AAAGACAAAG GTTTCGACTC TTTGACGCTG GATCAGTC  
     101 TCAGGAAAAA CGAGAAACTG AAGCTGGCGG CACAAGGTGC GGAAAAAAACT  
    151 TATGGAAACG GTGACAGCCT CAATACGGGC AAATTGAAGA ACGACAAGGT  
 201 CAGCCGTTTC GACTTTATCC GCCAAATCGA AGTGGACGGG CAGCTCATTA

10                  251 CCTTGGAGAG TGGAGAGTTC CAAGTATAACA AACAAAGCCA TTCCGCCTTA  
   301 ACCGCCTTTC AGACCGAGCA AATACAAGAT TCGGAGCATT CCGGGAAGAT  
   351 GGTTGCGAAA CGCCAGTTCA GAATCGGCCA CATAGCGGGC GAACATACAT  
   401 CTTTTGACAA GCTTCCCAGA GGCAGCAGGG CGACATATCG CGGGACGGCG  
   451 TTCGGTTCAG ACGATGCCGG CGGAAACTG ACCTACACCA TAGATTCGC  
   501 CGCCAAGCAG GGAAACGGCA AAATCGAACAA TTTGAAATCG CCAGAACTCA  
   551 ATGTCGACCT GGCGCCCGCC GATATCAAGC CGGATGGAAA ACGCCATGCC  
   601 GTCATCAGCG GTTCCGTCCT TTACAACCAA GCCGAGAAAG GCAGTTACTC  
   651 CCTCGGTATC TTTGGCGGAA AAGCCCAGGA AGTTGCCGGC AGCGCGGAAG  
   701 TGAAAACCGT AAACGGCATA CGCCATATCG GCCTTGCCGC CAAGCAACTC  
   751 GAGGGTGGCG GAGGCACTGG ATCCGCCAAC AACGACGACG ATGTTAAAAA  
   801 AGCTGCCACT GTGGCCATTG CTGCTGCCA CAACAAATGGC CAAGAAATCA  
   851 ACGGTTTCAA AGCTGGAGAG ACCATCTACG ACATTGATGA AGACGGCACA  
   901 ATTACCAAAA AAGACGCAAC TGCAGCCGAT GTTGAAGCCG AGCACTTTAA  
   951 AGGTCTGGGT CTGAAAAAAG TCGTGACTAA CCTGACCAAA ACCGTCATG  
 1001 AAAACAAACA AAACGTCATGCCAAAGTAA AAGCTGCAGA ATCTGAAATA  
 1051 GAAAAGTTAA CAACCAAGTT AGCAGACACT GATGCCGCTT TAGCAGATAC  
 1101 TGATGCCGCT CTGGATGCAA CCACCAACGC CTTGAATAAA TTGGGAGAAA  
 1151 ATATAACGAC ATTTGCTGAA GAGACTAAGA CAAATATCGT AAAAAATTGAT  
 1201 GAAAATTTAG AAGCCGTGGC TGATACCGTC GACAAGCATG CCGAAGCATT  
 1251 CAACGATATC GCCGATTATC TGGATGAAAC CAACACTAAG GCAGACGAAAG  
 1301 CCGTCAAAAC CGCCAATGAA GCCAAACAGA CGGGCGAAGA AACCAAACAA  
 1351 AACGTCGATG CCAAAGTAAAG AGCTGCAGAA ACTGCAGCAG GCAAAGCCGA  
 1401 AGCTGCCGCT GGCACAGCTA ATACTGCAGC CGACAAGGCC GAAGCTGTCG  
 1451 CTGAAAAGT TACCGACATC AAAGCTGATA TCGCTACGAA CAAAGATAAT  
 1501 ATTGCTAAAAA AAGCAAACAG TGCCGACGTG TACACCAGAG AAGAGTCTGA  
 1551 CAGCAAATTG GTCAGAATTG ATGGTCTGAA CGCTACTTACCA GAAAATTGG  
 1601 ACACACGCTT GGCTCTGCT GAAAATCCA TTGCCGATCA CGATACTCGC  
 1651 CTGAACGTT TGGATAAAAC AGTGTCAAGAC CTGCGCAAAG AAACCCGCCA  
 1701 AGGCCTTGCA GAACAAGCCG CGCTCTCCGG TCTGTTCCAA CCTTACAACG  
 1751 TGGGTCGTT CAATGTAACG GCTGCAGTCG CGGGCTACAA ATCCGAATCG  
 1801 GCAGTCGCCA TCGGTACCGG CTTCCGCTTT ACCGAAACT TTGCCGCCAA  
 1851 AGCAGGGCTG GCAGTCGGCA CTTCGTCCGG TTCTCCGCA GCCTACCATG  
 1901 TCGGCGTCAA TTACGAGTGG CTCGAGCACC ACCACCAACCA CCACTGA

40                  1 MVAADIGAGL ADALTAPLDH KDKGLQSLTL DQSVRKNEKL KLAQGAEKT  
   51 YGNQDSLNTG KLKNQKVSF DFIRQIEVDG QLITLESGEF QVYKQSHSAL  
   101 TAFQTEQIJD SEHSGKMKVAK RQFRIGDIAG EHTSFDFKLPE GGRATYRGTA  
   151 FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVDLAAA DIKPDKRHA  
   201 VISGSVLYNQ AEKGYSLSGI FGGKAQEVAQ SAEVKTVNGI RHIGLAAKQL  
   251 EGGGGTGSAT NDDDVKAAT VAIAAAYNNQ QEINGFKAGE TIYDIDEDEGT  
   301 ITKKDATAAD VEADDFKGLG LKKVVTNLTK TVNENKQNVD AKVKAASEEI  
   351 EKLTTKLADT DAALADTAA LDATTNALNK LGENDTTFAE ETKTNIVKID  
   401 EKLEAVADTV DKHAEAFNDI ADSLDETNTK ADEAVKTANE AKQTAEEETKQ  
   451 NVDAKVKAEE TAAGKAEEA GTANTAADKA EAVAAKVTDI KADIATNKN  
   501 IAKKANSADV YTREESDSKF VRIDGLNATT EKLDTRLASA EKSIADHDTR  
   551 LNGLDKTVSD LRKETRQGLA EQAALSGLFQ PYNVGRFNVT AAVGGYKSES  
   601 AVAIGTGFRF TENFAAKAGV AVGTSSGSSA AYHGVNYEW LEHHHHHH\*

AG741-961c

5                   1 ATGGTCGCCG CCGACATCGG TGCAGGGCTT GCCGATGCAC TAACCGCACC  
   51 GCTCGACCAT AAAGACAAAG GTTGACGCTG TTTGACGCTG GATCAGTCCG  
  101 TCAGGAAAAA CGAGAAACTG AAGCTGGCGG CACAAGGTGC GGAAAAAAACT  
  151 TATGAAAACG GTGACAGCCT CAATACGGGC AAATTGAAGA ACGACAAGGT  
  201 CAGCCGTTTC GACTTATCC GCCAAATCGA AGTGGACGGG CAGCTCATTA  
  251 CCTTGGAGAG TGGAGAGTTC CAAGTATACA AACAAAGCCA TTCCGCCTTA  
  301 ACCGCCTTTC AGACCGAGCA ATACAAAGAT TCGGAGCATT CCGGGAAAGAT  
  351 GGTTGCGAAA CGCCAGTTCA GAATCGGCCA CATAGCGGGC GAACATACAT  
  401 CTTTGACAA GCTTCCCAGA GGCAGCAGGG CGACATATCG CGGGACGGCG  
  451 TTCGGTTCAG ACGATGCCGG CGGAAACTG ACCTACACCA TAGATTTCGC  
  501 CGCCAAGCAG GGAAACGGCA AAATCGAACAA TTTGAAATCG CCAGAACTCA  
  551 ATGTCGACCT GGCGCCCGCC GATATCAAGC CGGATGGAAA ACGCCATGCC  
  601 GTCATCAGCG GTTCCGTCCT TTACAACCAA GCCGAGAAAG GCAGTTACTC  
  651 CCTCGGTATC TTTGGCGGAA AAGCCCAGGA AGTTGCCGGC AGCGCGGAAG  
  701 TGAAAACCGT AAACCGCATA CGCCATATCG GCCTTGCCTGC CAAGCAACTC  
  751 GAGGGTGGCG GAGGCACTGG ATCCGCCACA AACGACGACG ATGTTAAAAAA

20               801 AGCTGCCACT GTGGCCATTG CTGCTGCCTA CAACAATGGC CAAGAAATCA  
  851 ACGGTTTCAA AGCTGGAGAG ACCATCTACG ACATTGATGA AGACGGCACA  
  901 ATTACCAAAA AAGACGCAAC TGCAGCCGAT GTTGAAGCCG ACCACTTTAA  
  951 AGGTCTGGGT CTGAAAAAAAG TCGTGAACCAA CCTGACCAAA ACCGTCAATG  
  1001 AAAACAAACA AAACGTCGAT GCCAAAGTAA AAGCTGCAGA ATCTGAAATA  
  1051 GAAAAGTAA CAACCAAGTT AGCAGACACT GATGCCGCTT TAGCAGATAC  
  1101 TGATGCCGCT CTGGATGCAA CCACCAACGC CTTGAATAAA TTGGGAGAAA  
  1151 ATATAACGAC ATTTGCTGAA GAGACTAAGA CAAATATCGT AAAAATTGAT  
  1201 GAAAATTAG AAGCCGTGGC TGATACCGTC GACAAGCATG CCGAAGCATT  
  1251 CAACGATATC GCCGATTATCAT TGGATGAAAC CAAACACTAAG GCAGACGAAG  
  1301 CCGTCAAAAC CGCCAATGAA GCCAAACAGA CGGCCGAAGA AACCAAACAA  
  1351 AACGTCGATG CCAAAGTAA AGCTGCAGAA ACTGCAGCAG GCAAAGCCGA  
  1401 AGCTGCCGCT GGCACAGCTA ATACTGCAGC CGACAAGGCC GAAGCTGTCG  
  1451 CTGAAAAGT TACCGACATC AAAGCTGATA TCGCTACGAA CAAAGATAAT  
  1501 ATTGCTAAAA AAGCAAACAG TGCCGACGTG TACACCAGAG AAGAGTCTGA  
  1551 CAGCAAATTG GTCAGAATTG ATGGTCTGAA CGCTACTACC GAAAATTGG  
  1601 ACACACGCTT GGCTCTGCT GAAAATCCA TTGCCGATCA CGATACTCGC  
  1651 CTGAACGTT TGGATAAAAC AGTGTCAAGAC CTGCGCAAAG AAACCCGCCA  
  1701 AGGCCITGCA GAACAAGCCG CGCTCTCCGG TCTGTTCCAA CCTTACAACG  
  1751 TGGGTCTCGA GCACCACCA CACCACCACT GA

40               1 MVAADIGAGL ADALTAFLDH KDKGLQSLTL DQSVRKNEKL KLAAQGAEK  
  51 YNGNDSLNTG KLKNNDKVSF DFIRQIEVDG QLITLESGEF QVYKQSHSAL  
  101 TAFQTEQIJD SEHSGKMKVAK RQFRIGDIAG EHTSFDFKLPE GGRATYRGTA  
  151 FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA  
  201 VISGSVLYNQ AEKGSSYSLGI FGGKAQEVAQ SAEVKTVNGI RHIGLAAKQL  
  251 EGGGGTGSAT NDDDVKAAT VAIAAAYNNQ QEINGFKAGE TIYDIDEDEGT  
  301 ITKKDATAAD VEADDFKGLG LKKVVTNLTK TVNENKQNVD AKVKAEESEI  
  351 EKLTTKLADT DAALADTDAA LDATTNALNK LGENITTFAE ETKTNIVKID  
  401 EKLEAVADTV DKHAEAFNDI ADSLDETNTK ADEAVKTANE AKQTAEEETKO  
  451 NVDAKVAAE TAAGKAEAAA GTANTAADKA EAVAAKVTDI KADIATNKDN  
  501 IAIKANSADV YTREESDSKF VRIDGLNATT EKLDTRLASA EKSIADHDTR  
  551 LNGLDKTVSD LRKETRQGLA EQAALSGLFQ PYNVGLEHHH HHH\*

ΔG741-983

5	1	ATGGTCGCCG CCGACATCGG TGCGGGGCTT GCCGATGCAC TAACCGCACC
	51	GCTCGACCAT AAAGACAAAG GTTTGCAGTC TTTGACGCTG GATCAGTCCG
	101	TCAGGAAAAA CGAGAAAATG AAGCTGGCGG CACAAGGTGC GGAAAAAAACT
	151	TATGGAAACG GTGACAGCCT CAATACGGGC AAATTGAAGA ACGACAAGGT
	201	CAGCCGTTTC GACTTTATCC CCCAAATCGA AGTGGACGGG CAGCTCATTA
	251	CCTTGGAGAG TGGAGAGTTC CAAGTATAACA AACAAAGCCA TTCCGCCTTA
	301	ACCGCCTTTC AGACCGAGCA AATACAAGAT TC GGAGCATT CCGGGAAAGAT
	351	GGTTGCGAAA CGCCAGTTCA GAATCGGGCA CATAGCGGGC GAACATACAT
10	401	CTTTTGACAA GCTTCCCAGA GGCGGCAGGG CGACATATCG CGGGACGGCG
	451	TTCGGTTCAG ACGATGCCGG CGGAAAATCTG ACCTACACCA TAGATTTCGC
	501	CGCCAAGCGAG GGAAACGGCA AAATCGAACAA TTTGAAATCG CCAGAACTCA
	551	ATGTCGACCT GGCGCCCGCC GATATCAAGC CGGATGGAAA ACGCCATGCC
	601	GTCATCAGCG GTTCCGTCTT TTACAACCAA GCGAGAAAG GCAGTTACTC
15	651	CCTCGGTATC TTTGGCGGAA AAGCCCAGGA AGTTGCCGGC AGCGCGGAAG
	701	TGAAAACCGT AAACGGCATA CGCCATATCG GCCTTGCCGC CAAGCAACTC
	751	GAGGGATCCG GCGGAGGCAG CACTTCTGCG CCCGACTTCA ATGCAGGCGG
	801	TACCGGTATC GGCAGCAACA GCAGAGCAAC AACAGCGAAA TCAGCAGCAG
	851	TATCTTACCG CGGTATCAAG AACGAAATCTG GCAAAGACAG AAGCATGCTC
	901	TGTGCCGGTC GGGATGACGT TGCGGTTACA GACAGGGATG CCAAATCAA
20	951	TGCCCCCCCC CGGAATCTGC ATACCGGAGA CTTTCCAAC CCAAATGACG
	1001	CATACAAAGAA TTTGATCAAC CTCAAACCTG CAATTGAAGC AGGCTATACA
	1051	GGACCGGGGG TAGAGGTAGG TATCGTCGAC ACAGGCGAAT CGTCGGCAG
	1101	CATATCCTT CCCGAACTGT ATGGCAGAAA AGAACACGGC TATAACGAAA
	1151	ATTACAAAAA CTATACGGCG TATATGCGGA AGGAAGCGCC TGAAGACGGA
	1201	GGCGGTAAAG ACATTGAAGC TTCTTTCGAC GATGAGGCGG TTATAGAGAC
25	1251	TGAAGCAAAG CCGACGGATA TCCGCCACGT AAAAGAAATC GGACACATCG
	1301	ATTGGTCTC CCATATTATT GGCAGGGCTT CCGTGGACGG CAGACCTGCA
	1351	GGCGGTATTG CGCCCGATGC GACGCTACAC ATAATGAATA CGAATGATGA
	1401	AACCAAGAAC GAAATGATGG TTGCAGCCAT CCCAAATGCA TGGGTCAAGC
	1451	TGGGCGAACG TGGCGTGCAC ATCGTCAATA ACAGTTTGG ACAACATCG
	1501	AGGGCAGGCA CTGCCGACCT TTTCCAATA GCCAATTCCGG AGGAGCAGTA

30

35

40

45

50

55

5	1551	CCGCCAAGCG TTGCTCGACT ATTCCGGCG TGATAAAACA GACGAGGGTA
	1601	TCCGCCTGAT GCAACAGAGC GATTACGGCA ACCTGTCCTA CCACATCCGT
	1651	AATAAAAACA TGCTTTCAT CTTCGACAA GGCAATGACG CACAAGCTCA
	1701	GCCCCAACACA TATGCCCTAT TGCCCATTTA TGAAAAAGAC GCTCAAAAAG
	1751	GCATTATCAC AGTCGCAGGC GTAGACCGCA GTGGAGAAAA GTCAAAACGG
	1801	GAAATGTATG GAGAACCGGG TACAGAACCG CTTGAGTATG GCTCCAACCA
	1851	TTGCGGAATT ACTGCCATGT GGTGCCTGTC GGCACCTAT GAAGCAAGCG
	1901	TCCGTTTCAC CGGTACAAAC CCGATTCAA TTGCGGAAAC ATCCCTTTCC
	1951	GCACCCATCG TAACCGGCAC GGCAGCTCTG CTGCTGCAGA AATACCCGTG
10	2001	GATGAGCAAC GACAACCTGC GTACCACGTT GCTGACGACG GCTCAGGACA
	2051	TCGGTGCAGT CGGGCTGGAC AGCAAGTTCG GCTGGGGACT GCTGGATGCG
	2101	GGTAAGGCCA TGAACGGACC CGCGTCCTT CCGTTCGGCG ACTTTACCGC
	2151	CGATACGAAA GGTACATCCG ATATTGCCTA CTCCCTCCGT AACGACATTT
	2201	CAGGCACGGG CGGCCTGATC AAAAAAGGCG GCAGCCAAGT GCAACTGCAC
	2251	GGCAACACA CCTATACGGG CAAACCCATT ATCGAAGGCG GTTCGCTGGT
	2301	GTTGTACCGC AACACAACAAAT CGGATATGCG CGTCAAAAC AAAGGTGCGC
	2351	TGATTTATAA CGGGCGGC A TCCGGCGGC ACCTGAACAG CGACGGCATT
	2401	GTCTATCTGG CAGATACCGA CCAATCCGGC GCAAAACGAAA CCGTACACAT
	2451	CAAAGGCAGT CTGCAAGCTGG ACGGCAAAAGG TACGCTGTAC ACACGTTGG
	2501	GCAAACGTCT GAAAGTGGAC GGTACGGCA TTATCGGCG CAAGCTGTAC
	2551	ATGTCGGCAC GCGGCAAGGG GGCAGGCTAT CTCAACAGTA CCGGACGACG
20	2601	TGTTCCCTTC CTGAGTGGCG CCAAAATCGG GCAGGATTAT TCTTCTTCA
	2651	CAAACATCGA AACCGACGGC GGCCTGCTGG CTTCCCTCGA CAGCGTCGAA
	2701	AAAACAGCGG GCAGTGAAGG CGACACGCTG TCCTATTATG TCCGTCGC
	2751	CAATGCGCA CGGACTGCTT CGGCAGCGC ACATTCCGCG CCCGCCGGTC
	2801	TGAAACACGC CGTAGAACAG GGCAGCGAGCA ATCTGAAAAA CCTGATGGTC
	2851	GAACTGGATG CCTCCGAATC ATCCGCAACA CCCGAGACGG TTGAAACTGC
	2901	GGCAGCCGAC CGCACAGATA TCCGGCCAT CGCCCTTAC CGCGCAACTT
	2951	TCCGCGCAGC GGCAGCCGTA CAGCATGCGA ATGCGCCGA CGGTGTACGC
	3001	ATCTTCACA GTCTCGCGC TACCGTCTAT GCGCACAGTA CGGCCGCCCA
	3051	TGCGGATATG CAGGGACGCC GCCTGAAAGC CGTATCGGAC GGGTTGGACC
	3101	ACAACGGCAC GGGCTCTGCG GTCATCGCGC AAACCCAACA GGACGGTGG
30	3151	ACGTGGGAAC AGGGCGGTGT TGAAGGAAA ATGCGGGCA GTACCCAAAC
	3201	CGTCGGCATT GCCCGAAAAA CGGGCGAAAAA TACGACAGCA GCCGCCACAC
	3251	TGGGCATGGG ACGCAGCACA TGGAGCGAAA ACAGTGCAAA TGCAAAAC
	3301	GACAGCATTA GTCTGTTGC AGGCATACGG CACGATGCGG GCGATATCGG
	3351	CTATCTCAA GGCCTGTTCT CCTACGGACG CTACAAAAAC AGCATCAGCC
	3401	GCAGCAGCGG TGCAGACGAA CATCGGAAAG GCAGCGTCAA CGGCACGCTG
	3451	ATGCAAGCTGG GCGCACTGGG CGGTGTCAAC GTTCCGTTG CGCAACAGGG
	3501	AGATTGACG GTCGAAGGCG GTCTGCGCTA CGACCTGCTC AACAGGATG
	3551	CATTGCGCGA AAAAGGCAGT GCTTTGGGCT GGAGCGCAA CAGCCTCACT
	3601	GAAGGCACGC TGGTCGGACT CGCGGGCTG AAGCTGTCGC AACCTTGAG
	3651	CGATAAAAGCC GTCCGTGTTG CAACGGCGGG CGTGGAAACGC GACCTGAACG
	3701	GACGCGACTA CACGGTAACG GGCAGCTTTA CGGGCGCGAC TGCAGCAACC
40	3751	GGCAAGACGG GGGCACGAA TATGCCGAC ACCCGTCTGG TTGCCGGCCT
	3801	GGGCGCGGAT GTCGAATTG GCAACGGCTG GAACGGCTTG GCACGTTACA
	3851	GCTACGCCGG TTCCAAACAG TACGGCAACC ACAGCGGACG AGTCGGCGTA
	3901	GGCTACCGGT TCCTCGAGCA CCACCAACCA CACCACTGA

45

50

55

1 MVAADIGAGL ADALTAPLDH KDKGLQSLTL DQSVRKNEKL KLAAQGAEKT  
 51 YGNNGDSLNTG KLKNKDVSF DFIRQIEVDG QLITLESGEF QVYKQSHSAL  
 101 TAFQTEQIQD SEHSGKMKVAK RQFRIGDIAG EHTSFDFKLPE GGRATYRGTA  
 151 FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVDLAAA DIKPDGKRHA  
 201 VISGSVLYNQ AEKGSSYSLGI FGGKAQEVAQ SAEVKTVNGI RHIGLAAKOL  
 251 EGSGGGGTSA PDFNAGGTGI GSNSRATTAK SAAVSYAGIK NEMCKDRSML  
 301 CAGRDDVAVT DRDAKINAPP PNLHTGDFPN PNDAYKNLIN LKPAIEAGYT  
 351 GRGVVEVGIVD TGESVGSISF PELYGRKEHG YNENYKNYTA YMRKEAPEDG  
 401 GGKDIEASFD DEAVIETEAK PTDIRHVKEI GHIDLVSHII GGRSVDGRPA  
 451 GGIAPDATLH IMNTNDETKN EMMVAIRNA WVKLGERGVR IVNNNSFGTTS  
 501 RAGTADLFQI ANSEEQYRQA LLDYSGGDKT DEGIRLMOQS DYGNLSYHIR  
 551 NKNMLFIFST GNDAAQAPNT YALLPFYEKD AQKGIITVAG VDRSGEKFKR  
 601 EMYGEPGTEP LEYGSNHCGI TAMWCLSAPY EASVRFTRTN PIQIAGTSFS  
 651 APIVTGTAAL LLQKYPWMSN DNLRRTLLTT AQDIGAVGVD SKFGWGLDA  
 701 GKAMNGPASF PFGDFTADTK GTSDIAYSFR NDISGTGLI KKGGSQLQLH  
 751 GNNTYTGKTI IEGGSLVLYG NNKSDMRVET KGALIYNGAA SGGSLSNSDGI  
 801 VYLADTDQSG ANETVHIKGS LQLDGKGTLIY TRLGKLKVDT GATAIIGGKLY  
 851 MSARGKGAGY LNSTGRRVPF LSAAKIGQDY SFFTNIETDG GLLASLDSVE

20  
 901 KTAGSEGDTL SYVRRGNAA RTASAAAHSA PAGLKHAVEQ GGSNLENLMV  
 951 ELDASESSAT PETVETAAAD RTDMPGIRPY GATFRAAAAV QHANAADGVR  
 1001 IFNSLAATVY ADSTAHHADM QGRRRLKAVSD GLDHNGTGLR VIAQTQQDGG  
 1051 TWEQGGVEGK MRGSTQTVGI AAKTGENTTA AATLGMGRST WSENSANAKT  
 1101 DSISLFAGIR HDAGDIGYLN GLFSYGRYKN SISRSTGADE HAEGSVNGTL  
 1151 MQLGALGGVN VPFAATGDLT VEGGLRYDLL KQDAFAEKGS ALGWSGNSLT  
 1201 EGTLVGLAGL KLSQPLSDKA VLFATAGVER DLNGRDYTVT GGFTGATAAT  
 1251 GKTGARNMPH TRLVAGLGAD VEFNGNWGL ARYSYAGSKQ YGNHSGRVGV  
 1301 GYRFLEHHHH HH\*

30

35

40

45

50

55

ΔG741-ORF46.1

5	1	ATGGTCGCCG CCGACATCGG TGCGGGGCTT GCCGATGCAC TAACCGCACC
	51	GCTCGACCAT AAAGACAAGAAG GTTTGCGAGTC TTTGACGCTG GATCAGTCCG
	101	TCAGGAAAAA CGAGAAACTG AAGCTGGCGG CACAAGGTGC GGAAAAAAACT
	151	TATGGAACG GTGACAGCCT CAATACGGGC AAATTGAAGA ACGACAAGGT
	201	CAGCCGTTTC GACTTTATCC GCCAATTCGA AGTGGACGGG CAGCTCATT
	251	CCTTGGAGAG TGGAGAGTTC CAAGTATACA AACAAAGCCA TTCCGCCCTTA
	301	ACCGCCTTTC AGACCGAGCA AATAACAAGAT TCGGAGCATT CCGGGAAAGAT
	351	GGTTGCGAAA CGCCAGTTCA GAATCGGGCA CATAGCGGGC GAACATACAT
10	401	CTTTTGACAA GCTTCCCAGA GGCAGGAGGG CGACATATCG CGGGACGGCG
	451	TTCGGTTCAAG CAGATGCCGG CGGAAAAGT ACCTACACCA TAGATTCGCG
	501	CGCCAAGCAG GGAAACGGCA AAATCGAACAA TTTGAAATCG CCAGAACTCA
	551	ATGTCGACCT GGCGCCCGCC GATATCAAGC CGGATGGAAA ACGCCATGCC
	601	GTCATCAGCG GTTCCGTCCT TTACAACCAA GCGAGAAAG GCAGTTACTC
	651	CCTCGGTATC TTTGGCGGAA AAGCCAGGA AGTGGCCGGC AGCGCGGAAG
15	701	TGAAAACCGT AAACCGCATA CGCCATATCG GCCTTGCCGC CAAGCAACTC
	751	GACGGTGGCG GAGGCACTGG ATCCTCAGAT TTGGCAAACG ATTCTTTTAT
	801	CCGGCAGGTT CTCGACCGTC AGCATTTCGA ACCCGACGGG AAATACCACC
	851	TATTCCGAG CAGGGGGAA CTTGCCAGGC GCAGGGCCA TATCGGATTG
	901	GGAAAAATAC AAAGCCATCA GTTGGGCAAC CTGATGATTCA AACAGGCGGC
	951	CATTAAAGGA AATATCGGCT ACATTGTCCG CTTTCCGAT CACGGGCACG
20	1001	AAGTCCATTC CCCCTTCGAC AACATGCCT CACATCCGA TTCTGATGAA
	1051	GCCGGTAGTC CCGTTGACGG ATTTCAGCTT TACCGCATCC ATTGGGACGG
	1101	ATACGAACAC CATCCCGCC CGGGCTATGA CGGGCCACAG GGCGGCGGCT
	1151	ATCCCGCTCC CAAAGGCGCG AGGGATATAT ACAGCTACGA CATAAAAGGC
	1201	GTTGCCAAA ATATCCGCCT CAACCTGACC GACAACCGCA GCACCGGACA
	1251	ACGGCTTGC GACCGTTTCC ACAATGCCGG TAGTATGCTG ACGCAAGGAG
	1301	TAGGGCAGG ATTCAAACCG GCCACCCGAT ACAGCCCCGA GCTGGACAGA
	1351	TCGGGCAATG CGCCGAAAGC CTTCAACGGC ACTGCAGATA TCGTTAAAAA
	1401	CATCATCGGC GCGGCAGGAG AAATTGTCGG CGCAGGCGAT GCCGTGCAGG
	1451	GCATAAGCGA AGGCTCAAAC ATTGTCGTCA TGCACTGGCTT GGGTCTGCTT
	1501	TCCACCGAAA ACAAGATGGC GCGCATCAAC GATTTGGCAG ATATGGCGCA
30	1551	ACTCAAAGAC TATGCCGAG CAGCCATCCG CGATTGGCA GTCCAAAACC
	1601	CCAATGCCGC ACAAGGCATA GAAGCCGTCA GCAATATCTT TATGGCAGCC
	1651	ATCCCCATCA AAGGGATTGG AGCTGTTCGG GGAAAATACG GCTTGGGCGG
	1701	CATCACGGCA CATCTATCA AGCGGTCGCA GATGGCGCG ATCGCATTGC
	1751	CGAAAGGGAA ATCCGCCGTC AGCGACAATT TTGCCGATGC GGCAATACGCC
	1801	AAATACCGT CCCCTTACCA TTCCCGAAAT ATCCGTTCAA ACTTGGAGCA
35	1851	GCGTTACGGC AAAGAAAACA TCACCTCCTC AACCGTGCCT CGCTCAAACG
	1901	GCAAAATGT CAAACTGGC GACCAACGCC ACCCGAAGAC AGGCGTACCG
	1951	TTTGACGGTA AAGGGTTTCC GAATTTGAG AAGCACGTGA AATATGATAC
	2001	GCTCGAGCAC CACCAACCACC ACCACTGA
40	1	MVAADIGAGL ADALTAFLDH KDKGLQSLTL DQSVRKNEKL KLAAGQAEKT
	51	YGNQDSLNTG KLKNNDKVSRF DFIRQIEVDG QLITLESGEF QVYKQSHSAL
	101	TAFQTEQIQL SEHSGKVMVAK RQFRIGDIAG EHTSFDKLPE GGRATYRGTA
	151	FGSDDAGGKL TYTIDFAAKQ GNGKIEHLKS PELNVLAAA DIKPDGKRHA
	201	VISGSVLYNQ AEKGSSYSLGI FGGKAQEVEAG SAEVKTVNGI RHIGLAAKQL
	251	DGGGGTGSSD LANDSFIRQV LDRQHFEPDG KYHLFGSRGE LAERSGHIGL
	301	GKIQSHQLGN LMIQQAIIKG NIGYIVRFSD HGHEVHSPFD NHASHSDSDE
	351	AGSPVDFGSL YRIHWGDGYEH HPADGYDGPQ GGGYPAPKGA RDIYSYDIKG
	401	VAQNIRLNLT DNRSTGQRLL DRFHNAQGML TQGVGDGFKR ATRYSPELDR
	451	SGNAAEAFNG TADIVKNIIG AAGEIVGAGD AVQGISEGSN IAVMHGLLL
	501	STENKMARIN DLADMAQLKD YAAAAIRDWA VQNPNAAQGI EAVSNIFMAA
	551	IPIKGIGAVR GKYGIGGIT AHPIKRSQMGA IALPKGKSAV SDNFADAAYA
	601	KYPSPYHSRN IRSNLEQRYG KENITSSTVP PSNGKNVKLA DQRHPKTGVP
	651	FDGKGFPNFE KHVKYDTLEH HHHHH*

## 55 Example 16 - C-terminal fusions ('hybrids') with 287/ΔG287

[0228] According to the invention, hybrids of two proteins A & B may be either NH<sub>2</sub>-A-B-COOH or NH<sub>2</sub>-B-A-COOH. The effect of this difference was investigated using protein 287 either C-terminal (in '287-His' form) or N-terminal (in

$\Delta$ G287 form - sequences shown above) to 919, 953 and ORF46.1. A panel of strains was used, including homologous strain 2996. FCA was used as adjuvant:

	287 & 919		287 & 953		287 & ORF46.1	
Strain	$\Delta$ G287-919	919-287	$\Delta$ G287-953	953-287	$\Delta$ G287-46.1	46.1-287
<b>2996</b>	128000	16000	65536	8192	16384	8192
<b>BZ232</b>	256	128	128	<4	<4	<4
<b>1000</b>	2048	<4	<4	<4	<4	<4
<b>MC58</b>	8192	1024	16384	1024	512	128
<b>NGH38</b>	32000	2048	>2048	4096	16384	4096
<b>394/98</b>	4096	32	256	128	128	16
<b>MenA (F6124)</b>	32000	2048	>2048	32	8192	1024
<b>MenC (BZ133)</b>	64000	>8192	>8192	<16	8192	2048

20 Better bactericidal titres are generally seen with 287 at the N-terminus (in the  $\Delta$ G form)

[0229] When fused to protein 961 [ $\text{NH}_2$ - $\Delta$ G287-961-COOH - sequence shown above], the resulting protein is insoluble and must be denatured and renatured for purification. Following renaturation, around 50% of the protein was found to remain insoluble. The soluble and insoluble proteins were compared, and much better bactericidal titres were obtained with the soluble protein (FCA as adjuvant):

	2996	BZ232	MC58	NGH38	F6124	BZ133
Soluble	65536	128	4096	>2048	>2048	4096
Insoluble	8192	<4	<4	16	n.d.	n.d.

[0230] Titres with the insoluble form were, however, improved by using alum adjuvant instead:

35	Insoluble	32768	1128	4096	>2048	>2048	2048
----	-----------	-------	------	------	-------	-------	------

#### Example 17 — N-terminal fusions ('hybrids') to 287

[0231] Expression of protein 287 as full-length with a C-terminal His-tag, or without its leader peptide but with a C-terminal His-tag, gives fairly low expression levels. Better expression is achieved using a N-terminal GST-fusion.

[0232] As an alternative to using GST as an N-terminal fusion partner, 287 was placed at the C-terminus of protein 919 ('919-287'), of protein 953 ('953-287'), and of proteins ORF46.1 ('ORF46.1-287'). In both cases, the leader peptides were deleted, and the hybrids were direct in-frame fusions.

[0233] To generate the 953-287 hybrid, the leader peptides of the two proteins were omitted by designing the forward primer downstream from the leader of each sequence; the stop codon sequence was omitted in the 953 reverse primer but included in the 287 reverse primer. For the 953 gene, the 5' and the 3' primers used for amplification included a *Nde*I and a *Bam*H I restriction sites respectively, whereas for the amplification of the 287 gene the 5' and the 3' primers included a *Bam*H I and a *Xba*I restriction sites respectively. In this way a sequential directional cloning of the two genes in pET21b+, using *Nde*I-*Bam*H I (to clone the first gene) and subsequently *Bam*H I-*Xba*I (to clone the second gene) could be achieved.

[0234] The 919-287 hybrid was obtained by cloning the sequence coding for the mature portion of 287 into the *Xba*I site at the 3'-end of the 919-His clone in pET21b+. The primers used for amplification of the 287 gene were designed for introducing a *Sal*I restriction site at the 5'-and a *Xba*I site at the 3'- of the PCR fragment. Since the cohesive ends produced by the *Sal*I and *Xba*I restriction enzymes are compatible, the 287 PCR product digested with *Sal*I-*Xba*I could be inserted in the pET21 b-919 clone cleaved with *Xba*I.

[0235] The ORF46.1-287 hybrid was obtained similarly.

[0236] The bactericidal efficacy (homologous strain) of antibodies raised against the hybrid proteins was compared

with antibodies raised against simple mixtures of the component antigens:

	Mixture with 287	Hybrid with 287
919	32000	16000
953	8192	8192
ORF46.1	128	8192

[0237] Data for bactericidal activity against heterologous MenB strains and against serotypes A and C were also obtained for 919-287 and 953-287:

Strain	919		953		ORF46.1	
	Mixture	Hybrid	Mixture	Hybrid	Mixture	Hybrid
MC58	512	1024	512	1024	-	1024
NGH38	1024	2048	2048	4096	-	4096
BZ232	512	128	1024	16	-	-
MenA (F6124)	512	2048	2048	32	-	1024
MenC (C11)	>2048	n.d.	>2048	n.d.	-	n.d.
MenC (BZ133)	>4096	>8192	>4096	<16	-	2048

[0238] Hybrids of ORF46.1 and 919 were also constructed. Best results (four-fold higher titre) were achieved with 919 at the N-terminus.

[0239] Hybrids 919-519His, ORF97-225His and 225-ORF97His were also tested. These gave moderate ELISA titres and bactericidal antibody responses.

#### **Example 18 - the leader peptide from ORF4**

[0240] As shown above, the leader peptide of ORF4 can be fused to the mature sequence of other proteins (e.g. proteins 287 and 919). It is able to direct lipidation in *E.coli*.

#### **Example 19 - domains in 564**

[0241] The protein '564' is very large (2073aa), and it is difficult to clone and express it in complete form. To facilitate expression, the protein has been divided into four domains, as shown in figure 8 (according to the MC58 sequence):

Domain	A	B	C	D
Amino Acids	79-360	361-731	732-2044	2045-2073

[0242] These domains show the following homologies:

- Domain A shows homology to other bacterial toxins:

gbIAG03431.1|AE004443\_9probable hemagglutinin [Pseudomonas aeruginosa] (38%)  
 gbIAAC31981.1| (139897) HecA [Pectobacterium chrysanthemi] (45%)  
 emblCA36409.1| (X52156) filamentous hemagglutinin [Bordetella pertussis] (31%)  
 gbIAAC79757.1| (AF057695) large supernatant protein1 [Haemophilus ducreyi] (26%)  
 gbIAAA25657.1| (M30186) HpmA precursor [Proteus mirabilis] (29%)

- Domain B shows no homology, and is specific to 564.
- Domain C shows homology to:

gbIAAF84995.1|AE004032 HA-like secreted protein [Xylella fastidiosa] (33%)  
 gbIAAG05850.1|AE004673 hypothetical protein [Pseudomonas aeruginosa] (27%)  
 gbIAAF68414.1|AF237928 putative FHA [Pasteurella multocida] (23%)  
 5 gbIAAC79757.1|(AF057695) large supernatant protein1 [Haemophilus ducreyi] (23%)  
 pirlIS21010 FHA B precursor [Bordetella pertussis] (20%)

- Domain D shows homology to other bacterial toxins:

gbIAAF84995.1|AE004032\_14 HA-like secreted protein [Xylella fastidiosa] (29%)

10 [0243] Using the MC58 strain sequence, good intracellular expression of 564ab was obtained in the form of GST-fusions (no purification) and his-tagged protein; this domain-pair was also expressed as a lipoprotein, which showed moderate expression in the outer membrane/supernatant fraction.

15 [0244] The b domain showed moderate intracellular expression when expressed as a his-tagged product (no purification), and good expression as a GST-fusion.

[0245] The c domain showed good intracellular expression as a GST-fusion, but was insoluble. The d domain showed moderate intracellular expression as a his-tagged product (no purification). The cd protein domain-pair showed moderate intracellular expression (no purification) as a GST-fusion.

[0246] Good bactericidal assay titres were observed using the c domain and the bc pair.

#### *Example 20 - the 919 leader peptide*

[0247] The 20mer leader peptide from 919 is discussed in example 1 above:

MKKYLFRAAL YGIAAAILAA

25 [0248] As shown in example 1, deletion of this leader improves heterologous expression, as does substitution with the ORF4 leader peptide. The influence of the 919 leader on expression was investigated by fusing the coding sequence to the *PhoC* reporter gene from *Morganella morganii* [Thaller et al. (1994) Microbiology 140:1341-1350]. The construct was cloned in the pET21-b plasmid between the *Nde*I and *Xho*I sites (Figure 9):

30  
 1 MKKYLFRAAL YGIAAAILAA AIPAGNDATT KPDLYYLKNE QAIDSLKLLP  
 51 PPPEVGSIQF LNDQAMYEKG RMLRNTERGK QAQADADLAA GGVATAFSGA  
 101 FGYPITEKDS PELYKLLTNM IEDAGDLATR SAKEHYMRIR PFAFYGTETC  
 151 NTKDQKKLST NGSYPSGHTS IGWATALVLA EVNPANQDAI LERGYQLGQS  
 201 RVICGYHWQS DVDAARIVGS AAVATLHSDP AFQAQLAKAK QEFAQKSQK\*

35 [0249] The level of expression of *PhoC* from this plasmid is >200-fold lower than that found for the same construct but containing the native *PhoC* signal peptide. The same result was obtained even after substitution of the T7 promoter with the *E.coli* *Plac* promoter. This means that the influence of the 919 leader sequence on expression does not depend on the promoter used.

40 [0250] In order to investigate if the results observed were due to some peculiarity of the 919 signal peptide nucleotide sequence (secondary structure formation, sensitivity to RNAases, etc.) or to protein instability induced by the presence of this signal peptide, a number of mutants were generated. The approach used was a substitution of nucleotides of the 919 signal peptide sequence by cloning synthetic linkers containing degenerate codons. In this way, mutants were obtained with nucleotide and/or amino acid substitutions.

45 [0251] Two different linkers were used, designed to produce mutations in two different regions of the 919 signal peptide sequence, in the first 19 base pairs (L1) and between bases 20-36 (S1).

50 L1: 5' T ATG AAa/g TAc/t c/tTN TTt/c a/cGC GCC GCC CTG TAC GGC ATC GCC GCC  
 GCC ATC CTC GCC GCC GCG ATC CC 3'  
 S1: 5' T ATG AAA AAA TAC CTA TTC CGa/g GCN GCN c/tTa/g TAc/t GGc/g ATC GCC  
 GCC GCC ATC CTC GCC GCC GCG ATC CC 3'

55 [0252] The alignment of some of the mutants obtained is given below.

**L1 mutants:**

5           9L1-a   ATGAAGAAGTACCTTTCAGCGCCGCC~~~~~  
           9L1-e   ATGAAAAAAATACTTTTCCGCGCCGCC~~~~~  
           9L1-d   ATGAAAAAAATACTTTTCCGCGCCGCC~~~~~  
           9L1-f   ATGAAAAAAATATCTCTTAGCGCCGCCCTGTACGGCATCGCCGCCATCCTCGCCGCC  
           919sp   ATGAAAAAAATACCTATTCCGCGCCGCCCTGTACGGCATCGCCGCCATCCTCGCCGCC  
  
 10          9L1a   MKKYLFSAA~~~~~  
           9L1e   MKKYFFRAA~~~~~  
           9L1d   MKKYFFRAA~~~~~  
           9L1f   MKKYLFSAAALYGIAAAILAA  
           919sp   MKKYLFRALYGIAAAILAA (i.e. native signal peptide)

**S1 mutants:**

15          9S1-e   ATGAAAAAAATACCTATT.....ATCGCCGCCGCCATCCTCGCCGCC  
           9S1-c   ATGAAAAAAATACCTATTCCGAGCTGCCAATACGGCATCGCCGCCATCCTCGCCGCC  
           9S1-b   ATGAAAAAAATACCTATTCCGGGCGCCAATACGGCATCGCCGCCATCCTCGCCGCC  
           9S1-i   ATGAAAAAAATACCTATTCCGGGCGGCTTGACGGGATCGCCGCCATCCTCGCCGCC  
           919sp   ATGAAAAAAATACCTATTCCGCGCCGCCCTGTACGGCATCGCCGCCATCCTCGCCGCC  
  
 20          9S1e   MKKYL.....IAAAILAA  
           9S1c   MKKYLFRAAQYGIAAAILAA  
           9S1b   MKKYLFRAAQYGIAAAILAA  
           9S1i   MKKYLFRALYGIAAAILAA  
           919sp   MKKYLFRALYGIAAAILAA

**[0253]** As shown in the sequences alignments, most of the mutants analysed contain in-frame deletions which were unexpectedly produced by the host cells.

**[0254]** Selection of the mutants was performed by transforming *E. coli* BL21(DE3) cells with DNA prepared from a mixture of L1 and S1 mutated clones. Single transformants were screened for high PhoC activity by streaking them onto LB plates containing 100 µg/ml ampicillin, 50µg/ml methyl green, 1 mg/ml PDP (phenolphthaleindiphosphate). On this medium PhoC-producing cells become green (Figure 10).

**[0255]** A quantitative analysis of PhoC produced by these mutants was carried out in liquid medium using pNPP as a substrate for PhoC activity. The specific activities measured in cell extracts and supernatants of mutants grown in liquid medium for 0, 30, 90, 180 min. were:

**CELL EXTRACTS**

40          **[0256]**

	0	30	90	180
control	0,00	0,00	0,00	0,00
9phoC	1,11	1,11	3,33	4,44
9S1e	102,12	111,00	149,85	172,05
9L1a	206,46	111,00	94,35	83,25
9L1d	5,11	4,77	4,00	3,11
9L1f	27,75	94,35	82,14	36,63
9S1b	156,51	111,00	72,15	28,86
9S1c	72,15	33,30	21,09	14,43
9S1i	156,51	83,25	55,50	26,64
phoCwt	194,25	180,93	149,85	142,08

## SUPERNATANTS

[0257]

	0	30	90	1801
control	0,00	0,00	0,00	0,00
9phoC	0,33	0,00	0,00	0,00
9S1e	0,11	0,22	0,44	0,89
9L1a	4,88	5,99	5,99	7,22
9L1d	0,11	0,11	0,11	0,11
9L1f	0,11	0,22	0,11	0,11
9S1b	1,44	1,44	1,44	1,67
9S1c	0,44	0,78	0,56	0,67
9S1i	0,22	0,44	0,22	0,78
phoCwt	34,41	43,29	87,69	177,60

[0258] Some of the mutants produce high amounts of PhoC and in particular, mutant 9L1a can secrete PhoC in the culture medium. This is noteworthy since the signal peptide sequence of this mutant is only 9 amino acids long. This is the shortest signal peptide described to date.

**Example 21— C-terminal deletions of Maf-related proteins**

[0259] MafB-related proteins include 730, ORF46 and ORF29.

[0260] The 730 protein from MC58 has the following sequence:

1	V <u>KPLRRLLTML</u> LAACAVAAAAA <u>LIQPALAADL</u> AQDPFITDNA QRQHYEPGGK
51	YHLFGDPRGS VSDRTGKINV IQDYTHQMGN LLIQQANING TIGYHTRFSG
101	HGHEEHAPFD NHAADSASEE KGNVDEGFTV YRLNWECHEH HPADAYDGPK
151	GGNYPKPTGA RDEYTYHVNG TARSIKLNPT DTRSIRQRIS DNYSNLGSNF
201	SDRADEANRK MFEHNAKLDR WGNSMEEFING VAAGALNPFI SAGEALGIGD
251	ILYGTRYAID KAAMRNIAPL PAEGKFAVIG GLGSVAGFEK NTREAVDRWI
301	QENPNAAETV EAVFNAAAAA KVAKLAKAAK PGKAASVGD F ADSYKKKLAL
351	SDSARQLYQN AKYREALDIH YEDLIRRKTG GSSKFINGRE IDAVTN DALI
401	QAKRTISAID KPKNFLNQKN RKQIKATIEA ANQQGKRAEF WFKYGVHSQV
451	KSYIESKG <b>GI</b> VKTGLGD*

[0261] The leader peptide is underlined.

[0262] 730 shows similar features to ORF46 (see example 8 above):

- as for Orf46, the conservation of the 730 sequence among MenB, MenA and gonococcus is high (>80%) only for the N-terminal portion. The C-terminus, from ~340, is highly divergent.
- its predicted secondary structure contains a hydrophobic segment spanning the central region of the molecule (aa. 227-247).
- expression of the full-length gene in *E. coli* gives very low yields of protein. Expression from tagged or untagged constructs where the signal peptide sequence has been omitted has a toxic effect on the host cells. In other words, the presence of the full-length mature protein in the cytoplasm is highly toxic for the host cell while its translocation to the periplasm (mediated by the signal peptide) has no detectable effect on cell viability. This "intracellular toxicity" of 730 is particularly high since clones for expression of the leaderless 730 can only be obtained at very low frequency using a *recA* genetic background (*E. coli* strains: HB101 for cloning; HMS174(DE3) for expression).

[0263] To overcome this toxicity, a similar approach was used for 730 as described in example 8 for ORF46. Four C-terminal truncated forms were obtained, each of which is well expressed. All were obtained from intracellular expression of His-tagged leaderless 730.

[0264] Form A consists of the N-terminal hydrophilic region of the mature protein (aa. 28-226). This was purified as a soluble His-tagged product, having a higher-than-expected MW.

[0265] Form B extends to the end of the region conserved between serogroups (aa. 28-340). This was purified as an insoluble His-tagged product.

[0266] The C-terminal truncated forms named C1 and C2 were obtained after screening for clones expressing high levels of 730-His clones in strain HMS174(DE3). Briefly, the pET21b plasmid containing the His-tagged sequence coding for the full-length mature 730 protein was used to transform the *recA* strain HMS 174(DE3). Transformants were obtained at low frequency which showed two phenotypes: large colonies and very small colonies. Several large and small colonies were analysed for expression of the 730-His clone. Only cells from large colonies over-expressed a protein recognised by anti-730A antibodies. However the protein over-expressed in different clones showed differences in molecular mass. Sequencing of two of the clones revealed that in both cases integration of an *E. coli* IS sequence had occurred within the sequence coding for the C terminal region of 730. The two integration events have produced in-frame fusion with 1 additional codon in the case of C1, and 12 additional codons in the case of C2 (Figure 11). The resulting "mutant" forms of 730 have the following sequences:

**730-C1 (due to an IS1 insertion - figure 11A)**

1	MADLAQDPFI TDNAQRQHYE PGGKYHLFGD PRGSVSDRTG KINVIQDYTH
51	QMGNLLIQQAA NINGTIGYHT RFSGHGHEEH APFDNHAADS ASEEKGNVDE
101	GFTVYRLNWE GHEHHPADAY DGPKGNNYPK PTGARDEYTY HVNGTARSIK
151	LNPTDTRSIR QRISDNYSNL GSNFSDRADE ANRKMFEHNA KLDRWGNMSE
201	FINGVAAGAL NPFISAGEAL GIGDILYGTR YAIDKAAMRN IAPLPAEGKF
251	AVIGGLGSVA GFEKNTREAV DRWIQENPNA AETVEAVFNV AAAAKVAKLA
301	AAAKPGKAAV SGDFADSYKK KLALSDSARQ LYQNAKYREA LDIHYEDLIR
351	RKTDGSSKFI NGREIDAVTN DALIQAR*

[0267] The additional amino acid produced by the insertion is underlined.

**730-C2 (due to an IS5 insertion - Figure 11B)**

1	MADLAQDPFI TDNAQRQHYE PGGKYHLFGD PRGSVSDRTG KINVIQDYTH
51	QMGNLLIQQAA NINGTIGYHT RFSGHGHEEH APFDNHAADS ASEEKGNVDE
101	GFTVYRLNWE GHEHHPADAY DGPKGNNYPK PTGARDEYTY HVNGTARSIK
151	LNPTDTRSIR QRISDNYSNL GSNFSDRADE ANRKMFEHNA KLDRWGNMSE
201	FINGVAAGAL NPFISAGEAL GIGDILYGTR YAIDKAAMRN IAPLPAEGKF
251	AVIGGLGSVA GFEKNTREAV DRWIQENPNA AETVEAVFNV AAAAKVAKLA
301	AAAKPGKAAV SGDFADSYKK KLALSDSARQ LYQNAKYREA <u>LGKVRISEI</u>
351	<u>LLG</u> *

[0268] The additional amino acids produced by the insertion are underlined.

[0269] In conclusion, intracellular expression of the 730-C1 form gives very high level of protein and has no toxic effect on the host cells, whereas the presence of the native C-terminus is toxic. These data suggest that the "intracellular toxicity" of 730 is associated with the C-terminal 65 amino acids of the protein.

[0270] Equivalent truncation of ORF29 to the first 231 or 368 amino acids has been performed, using expression with or without the leader peptide (amino acids 1-26; deletion gives cytoplasmic expression) and with or without a His-tag.

#### Example 22 - domains in 961

[0271] As described in example 9 above, the GST-fusion of 961 was the best-expressed in *E.coli*. To improve expression, the protein was divided into domains (figure 12).

[0272] The domains of 961 were designed on the basis of YadA (an adhesin produced by *Yersinia* which has been demonstrated to be an adhesin localized on the bacterial surface that forms oligomers that generate surface projection [Hoiczyk et al. (2000) EMBO J 19:5989-99]) and are: leader peptide, head domain, coiled-coil region (stalk), and membrane anchor domain.

[0273] These domains were expressed with or without the leader peptide, and optionally fused either to C-terminal His-tag or to N-terminal GST. *E.coli* clones expressing different domains of 961 were analyzed by SDS-PAGE and

western blot for the production and localization of the expressed protein, from over-night (o/n) culture or after 3 hours induction with IPTG. The results were:

	Total lysate (Western Blot)	Periplasm (Western Blot)	Supernatant (Western Blot)	OMV SDS-PAGE
961 (o/n)	-	-	-	
961 (IPTG)	+/-	-	-	
961-L (o/n)	+	-	-	+
961-L (IPTG)	+	-	-	+
961c-L (o/n)	-	-	-	
961 c-L (IPTG)	+	+	+	
961Δ <sub>1</sub> -L (o/n)	-	-	-	
961Δ <sub>1</sub> -L (IPTG)	+	-	-	+

[0274] The results show that in *E. coli*:

- 961-L is highly expressed and localized on the outer membrane. By western blot analysis two specific bands have been detected: one at ~45kDa (the predicted molecular weight) and one at ~180kDa, indicating that 961-L can form oligomers. Additionally, these aggregates are more expressed in the over-night culture (without IPTG induction). OMV preparations of this clone were used to immunize mice and serum was obtained. Using overnight culture (predominantly by oligomeric form) the serum was bactericidal; the IPTG-induced culture (predominantly monomeric) was not bactericidal.
- 961Δ<sub>1</sub>-L (with a partial deletion in the anchor region) is highly expressed and localized on the outer membrane, but does not form oligomers;
- the 961c-L (without the anchor region) is produced in soluble form and exported in the supernatant.

[0275] Titres in ELISA and in the serum bactericidal assay using His-fusions were as follows:

	ELISA	Bactericidal
961a (aa 24-268)	24397	4096
961b (aa 269-405)	7763	64
961c-L	29770	8192
961c (2996)	30774	>65536
961c (MC58)	33437	16384
961d	26069	>65536

[0276] *E.coli* clones expressing different forms of 961 (961, 961-L, 961Δ<sub>1</sub>-L and 961c-L) were used to investigate if the 961 is an adhesin (c.f. YadA). An adhesion assay was performed using (a) the human epithelial cells and (b) *E.coli* clones after either over-night culture or three hours IPTG induction. 961-L grown over-night (961Δ<sub>1</sub>-L) and IPTG-induced 961c-L (the clones expressing protein on surface) adhere to human epithelial cells.

[0277] 961c was also used in hybrid proteins (see above). As 961 and its domain variants direct efficient expression, they are ideally suited as the N-terminal portion of a hybrid protein.

#### Example 23 — further hybrids

[0278] Further hybrid proteins of the invention are shown below (see also Figure 14). These are advantageous when compared to the individual proteins:

ORF46.1-741

5	1	ATGTCAGATT TGGCAAACGA TTCTTTTATC CGGCAGGTTC TCGACCGTCA
	51	GCATTTCGAA CCCGACGGGA AATACCACCT ATTCGGCAGC AGGGGGGAAC
	101	TTGCCGAGCG CAGCGGCCAT ATCGGATTGG GAAAATACA AAGCCATCAG
	151	TTGGGCAACC TGATGATTCA ACAGGGCGCC ATTAAGGAA ATATCGGCTA
	201	CATTGTCGGC TTTTCGATC ACAGGGCACGA AGTCCATTCC CCCTTCGACA
	251	ACCATGCCTC ACATTCGAT TCTGATGAAG CCGGTAGTCC CGTTGACGGA
	301	TTTAGCCTT ACCGCATCCA TTGGGACCGA TACGAACACC ATCCCAGCGA
	351	CGGCTATGAC GGGCCACAGG GCGGCGGCTA TCCCGCTCCC AAAGGCGCGA
10	401	GGGATATATA CAGCTACGAC ATAAAAGGGC TTGCCAAAAA TATCCGCTC
	451	AACCTGACCG ACAACCGCAG CACCGGACAA CGGCTGCCG ACCGTTCCA
	501	CAATGCCGGT AGTATGCTGA CGCAAGGAGT AGGCAGCGGA TTCAAACGCG
	551	CCACCCGATA CAGCCCCGAG CTGGACAGAT CGGGCAATGC CGCCGAAGCC
	601	TTCAACGGCA CTGAGATAT CGTTAAAAC ATCATCGGCG CGGCAGGAGA
	651	AATTGTCGGC GCAGGCCATG CCGTGCAGGG CATAAGCGAA GGCTCAAACA
15	701	TTGCTGTATC GCACCGCTTG GGTCTGCTT CCACCGAAAA CAAGATGGCG
	751	CGCATCAACG ATTGGCAGA TATGGCGCAA CTCAAAGACT ATGCCGAGC
	801	AGCCATCCGC GATTGGCAG TCCAAAACCC CAATGGCGCA CAAGGCATAG
	851	AAGCCGTCAG CAATATCTTT ATGGCAGCCA TCCCATCAA AGGGATTGGA
	901	GCTGTCGGG GAAAATACGG CTTGGCGGC ATCACGGCAC ATCCTATCAA
	951	GCGGTGCGAG ATGGGCGCGA TCGCATTGCC GAAAGGGAAA TCCGCCGTCA
20	1001	GCGACAATTT TGCCGATGCC GCATACGCCA AATACCGTC CCCTTACCAT
	1051	TCCCGAAATA TCCGTTCAAA CTTGGAGCAG CGTTACGGCA AAGAAAACAT
	1101	CACCTCCTCA ACCGTGCCGC CGTCAAACGG CAAAATGTC AACTGGCAG
	1151	ACCAACGCCA CCCGAAGACA GGCGTACCGT TTGACGGTAA AGGGTTCCG
	1201	AATTGAGA AGCACGTGAA ATATGATAACG GGATCCGGAG GGGGTGGTGT
	1251	CGCCGCCGAC ATCGGTGCGG GGCTTGCCGA TGCACTAACC GCACCGCTCG
25	1301	ACCATAAAGA CAAAGGTTTG CAGTCTTGA CGCTGATCA GTCCGTCAGG
	1351	AAAAACGAGA AACTGAAGCT GGCGGCACAA GGTGGGAAA AACTTATGG
	1401	AAACGGTGAC AGCCTCAATA CGGGCAAAAT GAAGAACGAC AAGGTAGCC
	1451	GTTTCGACTT TATCCGCCAA ATCGAAGTGG ACGGGCAGCT CATTACCTTG
	1501	GAGAGTGGAG AGTTCAGAACT ATACAAACAA AGCCATTCCG CCTTAACCGC
30	1551	CTTTCAAGACC GAGCAAATAC AAGATTCGGA GCATTCCGGG AAGATGGTTG
	1601	CGAAACGCCA GTTCAGAAATC GGCGACATAG CGGGCGAACAA TACATCTTT
35	1651	GACAAGCTTC CCGAAGGCAGG CAGGGCGACA TATCGCGGGA CGGCCTTCGG
	1701	TTCAGACGAT GCCGGCGGAA AACTGACCTA CACCATAGAT TTCCGCCA
	1751	AGCAGGGAAA CGGCAAAATC GAACATTGTA AATGCCAGA ACTCAATGTC
	1801	GACCTGGCCG CCGCCGATAT CAAGCCGGAT GGAAAACGCC ATGCCGTAT
	1851	CAGCGGTTCC GTCCCTTACA ACCAAGCCGA GAAAGGCAGT TACTCCCTCG
40	1901	GTATCTTGG CGGAAAAGCC CAGGAAGTTG CGGGCAGCGC GGAAGTGAAA
	1951	ACCGTAAACG GCATACGCCA TATCGCCCTT GCGCCAAGC AACTCGAGCA
	2001	CCACCACAC CACCACTGA
45	1	MSDLANDSFI RQVLDRQHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ
	51	LGNLMIQQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG
	101	FSLYRIHWDG YEHHPADGYD GPQGGGYPAP KGARDIYSYD IKGVAQNIRL
	151	NLTDNRSTGQ RLADRPHNAG SMLTQGVGDG FKRASTRYSPE LDRSGNAAEA
	201	FNGTADIVKN IIGAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA
	251	RINDLADMAQ LKDYYAAAIR DWAVQNPNAQ OGIEAVSNIF MAAIPIKGIG
50	301	AVRGKYGLGG ITAHPIKRSQ MGAIALPKGK SAVSDNFADA AYAKYPSPYH
	351	SRNIRSNLEQ RYGENITSS TVPPSNGKNV KLADQRHPKT GVPFDGKGFP
	401	NFEKHVKYDT GSGGGGVAAD IGAGLADALT APLDHDKGL QSLTLQSVR
	451	KNEKLKLAAQ GAEKTYGNND SLNTGKLND KVSRFDFIRQ IEVDGQLITL
	501	ESGEFQVYKQ SHSALTAQFT EQIQDSEHSG KMVAKRQFRI GDIAGEHTSF
	551	DKLPEGGRAT YRGTAFGSDD AGGKLTYTID FAAKQNGKI EHLKSPELNV
55	601	DLAAADIKPD GKRHAVISGS VLYNQAEGKS YSLGIFGGKA QEVAGSAEVK
	651	TVNGIRHIGL AAKOLEHHHH HH*

**ORF46.1-961**

5	1	ATGTCAGATT	TGGCAAACGA	TTCTTTTATC	CGGCAGGTTC	TCGACCGTCA
	51	GCATTCGAA	CCCGACGGGA	AATACCACCT	ATTCGGCAGC	AGGGGGGAAC
	101	TTGCCGAGCG	CAGCGGCCAT	ATCGGATTGG	GAAAATACA	AAGCCATCAG
	151	TTGGGCAACC	TGATGATTCA	ACAGGCGGCC	ATTAAAGGAA	ATATCGGCTA
	201	CATTGTCCGC	TTTCCGATC	ACGGGCACGA	AGTCATTCC	CCCTTCGACA
	251	ACCATGCCTC	ACATCCGAT	TCTGATGAAG	CCGGTAGTCC	CGTTGACGGA
	301	TTTAGCCTT	ACCGCATCCA	TTGGGACGGA	TACGAACACC	ATCCCGCCGA
	351	CGGCTATGAC	GGGCCACAGG	GCGGCCGCTA	TCCCCGTC	AAAGGCGCGA
10	401	GGGATATATA	CAGCTACGAC	ATAAAAGGCG	TTGCCCAAAA	TATCCGCTC
	451	AACCTGACCG	ACAACCGCAG	CACCGGACAA	CGGCTTCCG	ACCGTTTCCA
	501	CAATGCCGT	AGTATGCTGA	CGCAAGGAGT	AGGCAGCGGA	TTCAAACGCG
	551	CCACCCGATA	CAGCCCGAG	CTGGACAGAT	CGGGCAATGC	CGCCGAAGCC
	601	TTCAACGCA	CTGCAAGATAT	CGTTAAAAAC	ATCATCGCG	CGGCAGGAGA
	651	AATTGTCCGC	GCAGGCGATG	CCGTGCAGGG	CATAAGCGAA	GGCTAAACA
15	701	TTGCTGTCAT	GCACGGCTTG	GGTCTGCTT	CCACCGAAAA	CAAGATGGCG
	751	CGCATCAACG	ATTGGCAGA	TATGGCGCAA	CTCAAAGACT	ATGCCGCAGC
	801	AGCCATCCGC	GATTGGCAG	TCCAAAACCC	CAATGCCGCA	CAAGGCATAG
	851	AAGCGTCAG	CAATATCTT	ATGGCAGCCA	TCCCCATCAA	AGGGATTGGA
	901	GCTGTTCGGG	GAAAATACGG	CTTGGGCGGC	ATCACGGCAC	ATCCTATCAA
	951	GCGGTCGAG	ATGGGCGCGA	TCGCATTGCC	GAAAGGGAAA	TCCGCCGTCA
20	1001	GCGACAATT	TGCCGATGCG	GCATACGCCA	AATACCGTC	CCCTTACCAT
	1051	TCCCCAAATA	TCCGGTCAAA	CTTGGAGCAG	CGTTACGGCA	AAAGAAAACAT
	1101	CACCTCCTCA	ACCGTGCCGC	CGTCAAACGG	AAAAAATGTC	AAACTGGCAG
	1151	ACCAACGCCA	CCCGAAGACA	GGCGTACCGT	TTGACGGTAA	AGGGTTTCCG
	1201	AATTGAGA	AGCACGTGAA	ATATGATACG	GGATCCGGAG	GAGGAGGAGC
25	1251	CACAAACGAC	GACGATGTTA	AAAAAGCTGC	CACTGTGGCC	ATTGCTGCTG
	1301	CCTACAAACAA	TGGCCAAGAA	ATCAACGGTT	TCAAAGCTGG	AGAGACCATC
	1351	TACGACATTG	ATGAAGACGG	CACAATTACC	AAAAAAGACG	CAACTGCAGC
	1401	CGATGTTGAA	GCCGACGACT	TTAAAGGTCT	GGGTCTGAAA	AAAGTCGTGA
	1451	CTAACCTGAC	CAAACCGTC	AATGAAAACA	AACAAAACGT	CGATGCCAAA
30	1501	GTAAAAGCTG	CAGAATCTGA	AATAGAAAAG	TTAACAAACCA	AGTTAGCAGA
	1551	CACTGATGCC	GCTTTAGCAG	ATACTGATGC	CGCTCTGGAT	GCAACCACCA
	1601	ACGCCCTGAA	TAATTTGGGA	GAAAATATAA	CGACATTGTC	TGAAGAGACT
	1651	AAGACAAATA	TCGTTAAAAT	TGATGAAAAA	TTAGAAGCCG	TGGCTGATAC
	1701	CGTCGACAAG	CATGCCGAAG	CATTCAACGA	TATGCCGAT	TCATTGGATG
	1751	AAACCAACAC	TAAGGCAGAC	GAAGCCGTCA	AAACCGCCAA	TGAAGCCAAA
	1801	CAGACGGCCG	AAGAAACCAA	ACAAAACGTC	GATGCCAAAG	TAAGAGCTGC
	1851	AGAAACTGCA	GCAGGCAAAG	CCGAAGCTGC	CGCTGGCACA	GCTAATACTG
	1901	CAGCCGACAA	GGCGGAAGCT	GTCGCTGCAA	AAGTTACCGA	CATCAAAGCT
	1951	GATATCGCTA	CGAACAAAGA	TAATATTGCT	AAAAAAGCAA	ACAGTGCCGA
	2001	CGTGTACACC	AGAGAAGAGT	CTGACAGCAA	ATTGTCAGA	ATTGATGGTC
40						
	2051	TGAACGCTAC	TACCGAAAAA	TTGGCACAC	GCTTGGCTTC	TGCTGAAAAA
45	2101	TCCATTGCCG	ATCACGATAC	TCGCCTGAC	GGTTGGATA	AAACAGTGTG
	2151	AGACCTGCGC	AAAGAAACCC	GCCAAGGCC	TGCAGAACAA	GCCGCGCTCT
	2201	CCGGTCTGTT	CCAAACCTTAC	AACGTGGGTC	GGTTCAATGT	AACGGCTGCA
	2251	GTCGGCGGCT	ACAAATCCGA	ATCGGCAGTC	GCCATCGGTA	CCGGCTTCCG
	2301	CTTTACCGAA	AACTTGCCTG	CCAAAGCAGG	CGTGGCAGTC	GGCACTTCGT
	2351	CCGGTTCTTC	CGCAGCCTAC	CATGTCGGCG	TCAATTACGA	GTGGCTCGAG
	2401	CACCAACCACC	ACCACCACTG	A		

50

55

1 MSDLANDSFI RQVLDRQHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ  
 51 LGNLMIQQA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG  
 101 FSLYRIHWDG YEHHPADGYD GPQGGGYPAP KGARDIYSYD IKGVAQNIRL  
 151 NLTDNRSTGQ RLADRHNAG SMLTQGVGDG FKRASTRYSPE LDRSGNAAEA  
 201 FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAMHGL GLLSTENKMA  
 251 RINDLADMAQ LKDYAAAIR DWAVQNPNAQ QGIEAVSNIF MAAIPIKGIG  
 301 AVRGGKYGLGG ITAHPIKRSQ MGAIALPKGK SAVSDNFADA AYAKYPSPYH  
 351 SRNIRSNEQ RYGENITSS TVPPSNGKVN KLADQRHPKT GVPFDGKGFP  
 401 NFEKHVKYDT GSGGGGATND DVVKKAATVA IAAAYNNQE INGFKAGETI  
 451 YDIDEDGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV NENKQNVDAK  
 501 VKAAESEIEK LTTKLADTDA ALADTDAALD ATTNALNKLG ENITTFAEET  
 551 KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDETNKAD EAVKTANEAK  
 601 QTAEETKQNV DAKVKAETA AGKAEAAAGT ANTAADKAEA VAAKVTDIKA  
 651 DIATNKDNIA KKANSADVYT REESDSKFVR IDGLNATTEK LDTRLASAEK  
 701 SIADHDTRLN GLDKTVSDLR KETRQGLAEQ AALSGLFQPY NVGRFNVTAA  
 751 VGGYKSESAB AIGTGFRTE NFAAKAGVAV GTSSGSSAAY HVGVNYEWLE  
 801 HHHHHH\*

**ORF46.1-961c**

20 1 ATGTCAGATT TGGCAAACGA TTCTTTTATC CGGCAGGTTT CCGACCGTCA  
 51 GCATTTCGAA CCCGACGGGA AATACCACCT ATTCCGGCAGC AGGGGGGAAC  
 101 TTGGCGAGCG CAGCGGCCAT ATCGGATTGG GAAAAATACA AAGCCATCAG  
 151 TTGGGCAACC TGATGATTCA ACAGGGCGGC ATTAAAGGAA ATATCGGCTA  
 201 CATTGTCCGC TTTTCCGATC ACAGGGCACGA AGTCCATTCC CCCTTCGACA  
 251 ACCATGCTC ACATTCGAT TCTGATGAAG CCGTAGTCC CGTTGACGGA  
 301 TTTAGCCTT ACCGCATCCA TTGGGACGGA TACGAACACC ATCCCGCCGA  
 351 CGGCTATGAC GGGCACAGG GCGGGCGCTA TCCCCTGCTCC AAAGGCGCGA  
 401 GGGATATATA CAGCTACGAC ATAAAAGGCG TTGCCCCAAA TATCCGCCTC  
 451 AACCTGACCG ACAACCGCAG CACCGGACAA CGGCTTGCCG ACCGTTCCA  
 501 CAATGCCGGT AGTATGCTGA CGCAAGGAGT AGGCGACGGA TTCAAACGCG  
 551 CCACCCGATA CAGCCCCGAG CTGGACAGAT CGGGCATATGC CGCCGAAGCC  
 601 TTCAACGGCA CTGAGATAT CGTTAAAAAC ATCATCGCG CGGCAGGAGA  
 651 AATTGTCGGC GCAGCGATG CGTGCAGGG CATAAGCGAA GGCTAAACAA  
 701 TTGCTGTCAT GCACGGCTTG GGTCTGCTT CCACCGAAAAA CAAGATGGCG  
 751 CGCATCAACG ATTGGCAGA TATGGCGCAA CTCAAAGACT ATGCCGCAGC  
 801 AGCCATCCGC GATTGGCAG TCCAAAACCC CAATGCCGA CAAGGCATAG  
 851 AAGCGTCAG CAATATCTT ATGGCAGCCA TCCCCATCAA AGGGATTGGA  
 901 GCTGTTCCGG GAAAATACGG CTTGGGGCGG ATCACGGCAC ATCCTATCAA  
 951 GCGGTCGCAG ATGGGCGCGA TCGCATTGCC GAAAGGGAAA TCCGCGTCA  
 1001 GCGACAATTG TGCGATGCG GCATACGCC AATACCCGTC CCCTTACCAT  
 1051 TCCCGAAATA TCCGTTCAA CTTGGAGCAG CGTACGGCA AAGAAAACAT  
 1101 CACCTCCTCA ACCGTGCCGC CGTCAAACGG CAAAATGTC AACTGCGAG  
 1151 ACCAACGCCA CCCGAAGACA GGCACCGT TTGACCGTAA AGGGTTTCCG  
 1201 AATTTGAGA AGCACGTGA ATATGATACG GGATCCGGAG GAGGAGGAGC  
 1251 CACAAACGAC GACGATGTT AAAAGCTGC CACTGTGGCC ATTGCTGCTG  
 1301 CCTACAACAA TGGCAAGAA ATCAACGGTT TCAAAGCTGG AGAGACCATC  
 1351 TACGACATTG ATGAAGACGG CACAATTACG AAAAGACG CAACTGCAGC  
 1401 CGATGTTGAA GCCGACGACT TAAAGGTCT GGGTCTGAAA AAAGTCGTGA  
 1451 CTAACCTGAC CAAAACCGTC AATGAAAACA AACAAAACGT CGATGCCAAA  
 1501 GTAAAAGCTG CAGAATCTGA AATAGAAAAG TTAACAACCA AGTTAGCAGA  
 1551 CACTGATGCC GCTTAGCAG ATACTGATGC CGCTCTGGAT GCAACCACCA  
 1601 ACGCCTGAA TAAATTGGGA GAAAATATAA CGACATTGC TGAAGAGACT  
 1651 AAGACAAATA TCGTAAAAT TGATGAAAAA TTAGAAGCCG TGGCTGATAC  
 1701 CGTCGACAAG CATGCCGAAG CATTCAACGA TATCGCCGAT TCATTGGATG  
 1751 AAACCAACAC TAAGGCAGAC GAAGCCGTCA AAACCGCCAA TGAAGCCAAA  
 1801 CAGACGGCCG AAGAAACCAA ACAAACGTC GATGCCAAAG TAAAAGCTGC  
 1851 AGAAACTGCA GCAGGCAAAG CGCAAGCTGC CGCTGGCACA GCTAATACTG

**EP 1 790 660 A2**

1901 CAGCCGACAA GGCGGAAGCT GTCGCTGCAA AAGTTACCGA CATCAAAGCT  
1951 GATATCGCTA CGAACAAAGA TAATATTGCT AAAAAGCAA ACAGTGCCGA  
2001 CGTGTACACC AGAGAAGAGT CTGACAGCAA ATTTGTCTAGA ATTGATGGTC  
5 2051 TGAACGCTAC TACCGAAAAA TTGGACACAC GCTTGGCTTC TGCTGAAAAA  
2101 TCCATTGCCG ATCACGATAC TCGCCTGAAC GGTTGGATA AAACAGTGTC  
2151 AGACCTGCGC AAAGAAAACCC GCCAAGGCCT TGCAGAACAA GCCCGCGCTCT  
2201 CCGGTCTGTT CCAACCTTAC AACGTGGTC TCGAGGCACCA CCACCACAC  
2251 CACTGA

10

1 MSDLANDSFI RQVLDRQHFE PDGKYHLFGS RGELAERSGH IGLGKIQSHQ  
51 LGNLMIQQAA IKGNIGYIVR FSDHGHEVHS PFDNHASHSD SDEAGSPVDG  
101 FSLYRIHWDG YEHHPADGYD GPQGGGYPAP KGARDIYSYD IKGVAQNIRL  
15 151 NLTDNRSTGQ RLADRHNAG SMLTQGVGDG FKRASTRYSPE LDRSGNAAEA  
201 FNGTADIVKN IIGAAGEIVG AGDAVQGISE GSNIAVMHGL GLLSTENKMA  
251 RINDLADMAQ LKDYAAAIR DWAVQNPNAQ OGIEAVSNIF MAAIPIKGIG  
301 AVRKYGLGG ITAHPIKRSQ MGAIALPKGK SAVSDNFADA AYAKYPSPYH  
351 SRNIRSNEQ RYGKENITSS TVPPSNGKNV KLADQRHPKT GVPFDKGFP  
401 NFEKHVKYDT GSGGGGATND DDVKAATVA IAAAYNNGQE INGFKAGETI  
451 YDIDEDGTIT KKDATAADVE ADDFKGLGLK KVVTNLTKTV NENKQNVDAK  
501 VKAAESEIEK LTTKLADTDA ALADTDAAALD ATTNALNKLG ENITTFAEET  
551 KTNIVKIDEK LEAVADTVDK HAEAFNDIAD SLDENTKAD EAVKTANEAK  
601 QTAEETKQNV DAKVKAETA AGKAEAAAAGT ANTAADKAEA VAAKVTDIKA  
651 DIATNKDNIA KKANSADVYT REESDSKFVR IDGLNATTEK LDTRLASAEK  
701 SIADHDTRLN GLDKTVSDLR KETRQGLAEQ AALSGLFQPY NVGLEHHHHH  
25 751 H\*

30

35

40

45

50

55

**961-ORF46.1**

5	1	ATGGCCACAA	ACGACGACGA	TGTTAAAAAA	GCTGCCACTG	TGGCCATTGC
	51	TGCTGCCTAC	AACAATGGCC	AAGAAATCAA	CGGTTCAAA	GCTGGAGAGA
	101	CCATCTACGA	CATTGATGAA	GACGGCACAA	TTACCAAAAA	AGACGCAACT
	151	GCAGCCGATG	TTGAAGCCGA	CGACTTTAA	GGTCTGGGT	TGAAAAAAAGT
	201	CGTGACTAAC	CTGACCAAA	CCGTCATGA	AAACAAACAA	AACGTCGATG
	251	CCAAAGTAAA	AGCTGCAGAA	TCTGAAATAG	AAAAGTTAAC	AACCAAGTTA
	301	GCAGACACTG	ATGCCGCTT	AGCAGATACT	GATGCCGCTC	TGGATGCAAC
	351	CACCAACG	TTGAATAAAAT	TGGGAGAAAA	TATAACGACA	TTTGCTGAAG
10	401	AGACTAAGAC	AAATATCGTA	AAAATTGATG	AAAATTAGA	AGCCGTGGCT
	451	GATACCGTGC	ACAAGCATGC	CGAACGATT	AACGATATCG	CCGATTCAATT
	501	GGATGAAACC	AAACACTAAGG	CAGACGAAGC	CGTCAAACACC	GCCAATGAAG
	551	CCAAACAGAC	GGCCGAAGAA	ACCAAACAAA	ACGTCGATGC	CAAAGTAAAA
	601	GCTCAGAAA	CTGCAGCAGG	CAAAGCCGAA	GCTGCCGCTG	GCACAGCTAA
15	651	TACTGCAGCC	GACAAGGCCG	AAGCTGTCGC	TGCAAAAGTT	ACCGACATCA
	701	AAGCTGATAT	CGCTACGAAC	AAAGATAATA	TTGCTAAAAA	AGCAAAACAGT
	751	GCCGACGTGT	ACACCAGAGA	AGAGTCTGAC	AGCAAATTG	TCAGAATTG
	801	TGGTCTGAAC	GCTACTACCG	AAAATTGGA	CACACGCTTG	GCTTCTGCTG
	851	AAAAATCCAT	TGCCGATCAC	GATACTCGCC	TGAACGGTTT	GGATAAAAACA
	901	GTGTCAGACC	TGCCGAAAGA	AACCCGCCAA	GGCCTTGCAG	AACAAGCCGC
20	951	GCTCTCCGGT	CTGTTCCAAC	CTTACAAACGT	GGGTGGTTTC	AATGTAACGG
	1001	CTGCAGTCGG	CGGCTACAAA	TCCGAATCGG	CAGTCGCCAT	CGGTACCGGC
	1051	TTCCGTTTA	CCGAAAACCTT	TGCCGCCAAA	GCAGGGTGTG	CAGTCGGCAC
	1101	TTCGTCGGGT	TCTTCCGCAG	CCTACCATGT	CGGCGTCAAT	TACGAGTGGG
	1151	GATCCGGAGG	AGGAGGGATCA	GATTTGGCAA	ACGATTCTTT	TATCCGGCAG
	1201	GTTCTCGACC	GTCAGCATT	CGAACCCGAC	GGGAAATACC	ACCTATTCGG
25	1251	CAGCAGGGGG	GAACCTGCCG	AGCGCAGCGG	CCATATCGGA	TTGGGAAAAAA
	1301	TACAAAGCCA	TCAGTTGGC	AACCTGATGA	TTCAACAGGC	GGCCATTAAA
	1351	GGAAATATCG	GCTACATTGT	CCGTTTTCC	GATCACGGGC	ACGAAGTCCA
	1401	TTCCCCCTTC	GACAACCATG	CCTCACATT	CGATTCTGAT	GAAGCCGGTA
	1451	GTCCCCGTGA	CGGATTTAGC	CTTACCGCA	TCCATTGGGA	CGGATACGAA
30	1501	CACCATCCCG	CCGACGGCTA	TGACGGGGCCA	CAGGGGGGCC	GCTATCCC
	1551	TCCCCAAAGGC	GCGAGGGATA	TATACAGCTA	CGACATAAAA	GGCGTTGCC
	1601	AAAATATCCG	CCTCAACCTG	ACCGACAACC	GCAGCACCGG	ACAACGGCTT
	1651	GCCGACCGTT	TCCACAATGC	CGGTAGTATG	CTGACGCAAG	GAGTAGGGCA
	1701	CGGATTCAAA	CGGCCACCC	GATACAGCCC	CGAGCTGGAC	AGATCGGGCA
	1751	ATGCCGCCGA	AGCCTTCAAC	GGCACTGCAG	ATATCGTTAA	AAACATCATC
35	1801	GGCGCGGAG	GAGAAATTGT	CGGCGCAGGC	GATGCCGTG	AGGGCATAAG
	1851	CGAAGGCTCA	AACATTGCTG	TCATGCACGG	CTTGGGTCTG	CTTCCACCG
	1901	AAAACAAGAT	GGCGCGCATC	AACGATTG	CAGATATGGC	GCAACTCAAA

40	1951	GAATATGCCG	CAGCAGCCAT	CCGGGATTGG	GCAGTCCAAA	ACCCCAATGC
	2001	CGCACAAAGGC	ATAGAAGCCG	TCAGCAATAT	CTTTATGGCA	GCCATCCCCA
	2051	TCAAAGGGAT	TGGAGCTGTT	CGGGGAAAT	ACGGCTTGGG	CGGCATCACG
	2101	GCACATCCTA	TCAAGCGGTC	GCAGATGGGC	GCGATCGCAT	TGCCGAAAGG
	2151	GAAATCCGCC	GTCAGCGACA	ATTTTGC	TGCGGCATAC	GCCAAATACC
45	2201	CGTCCCCTTA	CCATTCCCGA	AAATATCGTT	CAAACCTGG	GCAGCGTTAC
	2251	GGCAAAGAAA	ACATCACCTC	CTCAACCGTG	CCGCGCTAA	ACGGCAAAAA
	2301	TGTCAAAC	GCAGACCAAC	GCCACCCGAA	GACAGGC	CCGTTTGACG
	2351	GTAAAGGGTT	TCCGAATT	GAGAAAGCAGC	TGAAATATGA	TACGCTCGAG
	2401	CACCACCA	ACCACCACTG	A		

50

55

1 MATNDDDVKK AATVAAIAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
 51 AADVEADDK GLGLKKVVTN LTAKTVNENKQ NVDAKVAAE SEIEKLTTKL  
 101 ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA  
 151 DTVDKHAEAF NDIADSDET NTKADEAVKT ANEAKQTAEE TKQNVDAKVK  
 201 AAETAAGKAE AAAGTANTAA DKAEEAVAALK TDIKADIATN KDNIAKKANS  
 251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDCK  
 301 VSDLRKETRQ GLAEQAALSG LFQPYNVGRF NVTAAVGGYK SESAVAIGTG  
 351 FRFTENFAAK AGVAVGTSSG SSAAYHVGVN YEWGSGGGGS DLANDSFIRQ  
 401 VLDRQHFEPD GKYLHFGSRG ELAERSGHIG LGKIQSHQLG NLMIQQAAIK  
 451 GNIGYIVRFS DHGHEVHSPF DNHASHSDSD EAGSPVDGFS LYRIHWGDYE  
 501 HHPADGYDGP QGGGYPAPKG ARDIYSYDIK GVAQNIRNLN TDNRSTGQRL  
 551 ADRFHAGSM LTQGVGDGFK RATRYSPELD RSGNAAEAFN GTADIVKNII  
 601 GAAGEIVGAG DAVQGISEGS NIAVMHGLGL LSTENKMARI NDLADMAQLK  
 651 DYAAAIRDW AVQNPNAAQG IEAVSNIFMA AIPIKGIGAV RGKYGLGGIT  
 701 AHPIKRSQLMG AIALPKGKSA VSDNFADAAY AKYPSPYHSR NIRSNLEQRY  
 751 GKENITSSTV PPSNGKNVKL ADQRHPKTGV PFDKGFPNF EKHVKYDTLE  
 801 HHHHHH\*

**961-741**

20 1 ATGGCCACAA ACGACGACGA TGTAAAAAAA GCTGCCACTG TG GCCATTGC  
 51 TGCTGCCTAC AACATGGCC AAGAAATCAA CGGTTCAAA GCTGGAGAGA  
 101 CCATCTACGA CATTGATGAA GACGGCACAA TTACCAAAAA AGACGCAACT  
 151 GCAGCCGATG TTGAAGCCGA CGACTTTAA GGTCTGGTC TGAAAAAAAGT  
 201 CGTGACTAAC CTGACCAAAA CCGTCAATGA AAACAAACAA AACGTCGATG  
 251 CCAAAGTAAA AGCTGCAGAA TCTGAAATAG AAAAGTTAAC ACCAAGTTA  
 301 GCAGACACTG ATGCCGCTT AGCAGATACT GATGCCGTC TG GATGCAAC  
 351 CACCAACGCC TTGAATAAAT TGGGAGAAAA TATAACGACA TTTGCTGAAG  
 401 AGACTAAGAC AAATATCGTA AAAATTGATG AAAAATTAGA AGCCGTGGCT  
 451 GATACCGTCG ACAAGCATGC CGAACGATTC AACGATATCG CCGATTCAATT  
 501 GGATGAAACC AACACTAAGG CAGACGAAGC CGTCAAAACC GCCAATGAAAG  
 551 CCAAACAGAC GGCGGAAGAA ACCAAACAA ACGTCGATGC CAAAGTAAAA  
 601 GCTGCAGAAA CTGCAGCAGG CAAAGCCGAA GCTGCCGTC GCACAGCTAA  
 651 TACTGCAGCC GACAAGGCCG AAGCTGTCG TGCAAAAGTT ACCGACATCA  
 701 AAGCTGATAT CGCTACGAAC AAAGATAATA TTGCTAAAAA AGCAAAACAGT  
 751 GCCGACGTGT ACACCAGAGA AGAGTCTGAC AGCAAAATTG TCAGAATTG  
 801 TGGTCTGAAC GCTACTACCG AAAAATTGGA CACACGCTTG GCTTCTGCTG  
 851 AAAAATCCAT TGCGGATCAC GATACTGCC TGAAACGGTTT GGATAAAACAA  
 901 GTGTCA GACC TGCGAAAGA AACCCGCCAA GGCCTTGCAG AACAGCCGC  
 951 GCTCTCCGGT CTGTTCCAAC CTTACAACTG GGGTCGGTT AATGTAACGG  
 1001 CTGCAGTCGG CGGCTACAAA TCCGAATCGG CAGTCGCCAT CGGTACCGGC  
 1051 TTCCGTTTA CCGAAAACCT TGCCGCCAAA CGAGCGTGG CAGTCGGCAC  
 1101 TTGCTCCGGT CCTTCCGAG CCTACCATGT CGCGTCAAT TACGAGTGGG  
 1151 GATCCGGAGG GGGTGGTGTG CCGCCGACA TCGGTGCGGG GCTTGCCTG  
 1201 GCACTAACCG CACCGCTCGA CTATAAACGAC AAAGTTTGCA AGTCTTTGAC  
 1251 GCTGGATCG TCCGTCAGGA AAAACGAGAA ACTGAAGCTG GCGGCACAAG  
 1301 GTGCGAAAAA AACTTATGGA AACGGTGACA GCCTCAATAC GGGCAAATTG  
 1351 AAGAACGACA AGGTCA GCGC TTTCGACTTT ATCCGCCAA TCGAAGTGG  
 1401 CGGGCAGCTC ATTACCTTG AGAGTGGAGA GTTCCAAGTA TACAAACAAA  
 1451 GCCATTCCGC CTTAACCGCC TTTCAGACCG AGCAAATACA AGATTGGAG  
 1501 CATTCCGGGA AGATGGTTGC GAAACGCCAG TTCAGAATCG GCGACATAGC  
 1551 GGGCGAACAT ACATTTTG ACAAGCTTCC CGAACGGCGC AGGGCGACAT  
 1601 ATCGCGGGAC GGCCTTCGGT TCAGACGATG CGGGCGGGAA ACTGACCTAC  
 1651 ACCATAGATT TCGCGCAGCA GCAGGGAAAC GGCAAATCG AACATTGAA  
 1701 ATCGCCAGAA CTCAATGTCG ACCTGGCCGC CGCCGATATC AAGCCGGATG  
 1751 GAAAACGCCA TGCGCTCATC AGCGGTTCCG TCCTTACAA CCAAGCCGAG  
 1801 AAAGGCAGTT ACTCCCTCGG TATCTTGGC GGAAAAGCCC AGGAAGTTGC

55 1851 CGGCAGCGCG GAAGTGAAAAA CCGTAAACGG CATA CGCCAT ATCGGCCTTG  
 1901 CCGCCAAGCA ACTCGAGCAC CACCACCA ACCACTGA

**EP 1 790 660 A2**

1 MATNDDDVKK AATVAIAAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
51 AADVEADDK GLGLKKVVTN LTKTVNENKQ NVDAKVAAE SEIEKLTTKL  
101 ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA  
151 DTVDKHAEAF NDIADSDET NTKADEAVKT ANEAKQTAEE TKQNVDAKVK  
201 AAETAAGKAE AAAGTANTAA DKAEEAVAALK TDIKADIATN KDNIAKKANS  
251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT  
301 VSDLRKETRQ GLAEQAALSG LFQPYNVGRF NVTAAVGGYK SESAVAIGTG  
351 FRFTENFAAK AGVAVGTSSG SSAAYHVGVN YEWGSGGGGV AADIGAGLAD  
401 ALTAPLDHKD KGLQSLTLDO SVRKNEKLKL AAQGAEKTYG NGDSLNTGKL  
451 KNDKVSRFDF IRQIEVDGQL ITLESGEFQV YKQSHSALTA FQTEQIQDSE  
501 HSGKMVAKRQ FRIGDIAGEH TSFDKLPEGG RATYRGTAFG SDDAGGKLTY  
551 TIDFAAKQGN GKIEHLKSPE LNVDLAAADI KPDGKRHAVI SGSVLYNQAE  
601 KGSYSLGIFG GKAQEVAAGSA EVKTVNGIRH IGLAAKQLEH HHHHH\*

15

20

25

30

35

40

45

50

55

961-983

5	1	ATGGCCACAA ACGACGACGA TGTTAAAAAA GCTGCCACTG TGGCCATTGC
	51	TGCTGCCTAC AACAAATGGCC AAGAAATCAA CGGTTCAAA GCTGGAGAGA
	101	CCATCTACGA CATTGATGAA GACGGCACAA TTACCAAAAA AGACGCAACT
	151	GCAGCCGATG TTGAAGCCGA CGACTTAAAG GGTCTGGTC TGAAAAAAAGT
	201	CGTGACTAAC CTGACCAAAA CCGTCATGA AAACAAACAA AACGTCGATG
	251	CCAAAGTAAA AGCTGCAGAA TCTGAAATAG AAAAGTTAAC ACCAAGTTA
	301	GCAGACACTG ATGCGCTTT AGCAGATACT GATGCCGCTC TGGATGCAAC
	351	CACCAAGCC TTGAATAAT TGGGAGAAAA TATAACGACA TTGCTGAAG
10	401	AGACTAAGAC AAATATCGT AAAATTGATG AAAATTAGA AGCCGTGGCT
	451	GATACCGTCG ACAAGCATGC CGAACGATTC AACGATATCG CCGATTCAATT
	501	GGATGAAACC AACACTAAGG CAGACGAAGC CGTAAAACC GCACATGAAG
	551	CCAAACAGAC GGCGAAGAA ACCAAACAAA ACCTCGATGC CAAAGTAAA
	601	GCTGCAGAAA CTGCAGCAGG CAAAGCCGAA GCTGCCGCTG GCACAGCTAA
	651	TACTGCAGCC GACAAGGCCG AAGCTGTCGC TGCAAAAGTT ACCGACATCA
	701	AAGCTGATAT CGCTACGAAC AAAGATAATA TTGCTAAAAA AGCAAAACAGT
	751	GCCGACGTGT ACACCAGAGA AGAGTCTGAC AGCAAATTG TCAGAATTGA
	801	TGGTCTGAAC GCTACTACCG AAAAATTGGA CACACGCTTG GCTTCTGCTG
	851	AAAAATCCAT TGCCGATCAC GATACTCGCC TGAACGGTTT GGATAAAAACA
	901	GTGTCAAGGC ACCCCGAAAGA AACCCGCCAA GGCTTGCAG ACAAGCCGC
	951	GCTCTCCGGT CTGTCACCAAC CCTACACCGT GGGTCCGGTTC AATGTAACGG
20	1001	CTGCAGTCGG CGGCTACAAA TCCGAATCGG CAGTCGCCAT CGGTACCGGC
	1051	TTCCGTTTA CCGAAAACCTT TGCCGCCAAA GCAGGCGTGG CAGTCGGCAC
	1101	TTCGTCGGGT TCTTCGCAG CCTACCATGT CGGCCTCAAT TACGAGTGGG
	1151	GATCCGGCGG AGGGCGCACT TCTGCGCCCC ACTTCAATGC AGGCGGTACC
	1201	GGTATCGCA GCAACAGCAG AGCAACAAACA GCGAAATCAG CAGCAGTATC
	1251	TTACGCCGT ATCAAGAACG AAATGTGCAA AGACAGAACG ATGCTCTGTG
	1301	CCGGTCGGGA TGACGTTGCG GTTACAGACA GGGATGCCAA AATCAATGCC
	1351	CCCCCCCGA ATCTGCATAC CGGAGACTTT CCAAACCCAA ATGACGCATA
	1401	CAAGAATTG ATCAACCTCA AACCTGCAAT TGAAGCAGGC TATACAGGAC
	1451	GCGGGGTAGA GGTAGGTATC GTCGACACAG GCGAATCCGT CGGCAGCATA
	1501	TCCTTCCCG AACTGTATGG CAGAAAAGAA CACGGCTATA ACGAAAATTA
30	1551	CAAAAACCTAT ACGGCGTATA TGCGGAAGGA AGCGCTGAA GACGGAGGCG
	1601	GTAAAGACAT TGAAGCTTCT TTGACGATG AGGCCGTTAT AGAGACTGAA
	1651	GCAAAGCCGA CGGATATCCG CCACGTAAA GAAATCGGAC ACATCGATTT
	1701	GGTCTCCCAT ATTATTGGCG GGCCTCCGT GGACGGCAGA CCTGCAGGCG
	1751	GTATTGCGCC CGATGCGACG CTACACATAA TGAATACGAA TGATGAAACC
	1801	AAGAACGAAA TGATGGTTGC AGCCATCCGC AATGCGATGG TCAAGCTGGG
	1851	CGAACGTGGC GTGCGCATCG TCAATAACAG TTTTGGAAACA ACATCGAGGG
	1901	CAGGCACTGC CGACCTTTT CAAATAGCCA ATTCCGGAGGA GCAGTACCGC
	1951	CAAGCGTTGC TCGACTATTC CGGCGGTGAT AAAACAGACG AGGGTATCCG
	2001	CCTGATGCAA CAGAGCGATT ACGGCAACCT GTCCTACAC ATCCGTAATA
	2051	AAAACATGCT TTTCATCTTT TCGACAGGGCA ATGACGCACA AGCTCAGCCC
	2101	AAACACATATG CCCTATTGCC ATTGTTATGAA AAAGACGCTC AAAAAAGGCAT
	2151	TATCACAGTC GCAGGCGTAG ACCGCAGTGG AGAAAAGTTT AAAACGGAAA
	2201	TGTATGGAGA ACCGGGTACA GAACCGCTTG AGTATGGCTC CAACCATTC
	2251	GGAATTACTG CCATGTGGTG CCTGTCGGCA CCCTATGAAG CAAGCGTCCG
	2301	TTTCACCCGT ACAAAACCGA TTCAAATTGC CGGAACATCC TTTTCCGCAC
	2351	CCATCGTAAC CGGCACGGCG GCTCTGCTGC TGCAAGAAATA CCCGTGGATG

45

50

55

	2401	AGCAACGACA ACCTGGTAC CACGGTGTG ACGACGGCTC AGGACATCGG
5	2451	TGCAGTCGGC GTGGACAGCA AGTTGGTGTG GGGACTGCTG GATGCGGGTA
	2501	AGGCCATGAA CGGACCCGCG TCCTTCCGT TCGGGACTT TACCGCCGAT
	2551	ACGAAAGGTA CATCCGATAT TGCCTACTCC TTCCGTAACG ACATTCAGG
	2601	CACGGGCGGC CTGATCAAAA AAGGCGGCAG CCAACTGCAA CTGCACGGCA
	2651	ACAACACCTA TACGGGCAAA ACCATTATCG AAGGCGGTTC GCTGGTGTG
10	2701	TACGGCAACA ACAAAATCGGA TATGCGCGTC GAAACCAAAG GTGCGCTGAT
	2751	TTATAACGGG GCGGCATCCG GCGGCAGCCT GAACAGCGAC GGCATTGTCT
	2801	ATCTGGCAGA TACCGACCAA TCCGGCGCAA ACGAAACCGT ACACATCAA
	2851	GGCAGTCGAGCTGGACGG CAAAGGTACG CTGTACACAC GTTTGGGCAA
	2901	ACTGCTGAAA GTGGACGGTA CGCGGATTAT CGGCACGAAAG CTGTACATGT
	2951	CGGCACGCGG CAAGGGGGCA GGCTATCTCA ACAGTACCGG ACCGACGTGTT
	3001	CCCTTCCTGA GTGCCGCCAA ATCGGGCAG GATTATTCTT TCTTCACAAA
15	3051	CATCGAAACC GACGGCGGCC TGCTGGCTTC CCTCGACAGC GTCGAAAAAA
	3101	CAGCGGGCAG TGAAGGCGAC ACGCTGTCT ATTATGTCCG TCGCGGCAAT
	3151	GCGGCACCGA CTGCTTCGGC AGCGGCACAT TCCGCGCCCG CCGGTCTGAA
	3201	ACACGCCGTA GAACAGGGCG GCAGCAATCT GGAAAACCTG ATGGTCGAAC
	3251	TGGATGCCCT CGAACATCC GCAACACCCC AGACGTTGA AACTGCGGCA
	3301	GCCGACCGCA CAGATATGCC GGGCATCCGC CCCTACGGCG CAACTTCCG
	3351	CGCAGCGCA GCCGTACAGC ATGCAATGC CGCCGACGGT GTACGCATCT
20	3401	TCAACAGTCT CGCCGCTACC GTCTATGCCG ACAGTACCGC CGCCCATGCC
	3451	GATATGCAGG GACGCCGCCT GAAAGCCGTA TCGGACGGGT TGGAACACAA
	3501	CGGCACGGGT CTGCGCGTCA TCGCGCAAAC CCAACAGGAC GTGGAACAGT
	3551	GGGAACAGGG CGGTGTTGAA GGCAAAATGCG GCGGCAGTAC CCAAACCGTC
	3601	GGCATTGCCG CGAAAACCGG CGAAAATACG ACAGCAGCCG CCACACTGGG
	3651	CATGGGACGC AGCACATGGA CGAAAACAG TGCAAATGCA AAAACCGACA
25	3701	GCATTAGTCT GTTTCGAGGC ATACGGCACCG ATGCGGGCGA TATCGGCTAT
	3751	CTCAAAGGCC TGTTCTCTA CGGACGCTAC AAAAACAGCA TCAGCCGAG
	3801	CACCGGTGCG GACGAACATG CGGAAGGGCAG CGTCAACGGC ACGCTGATGC
	3851	AGCTGGCGC ACTGGCGGT GTCAACGTT CGTTGCCGC AACGGGAGAT
	3901	TTGACGGTCA AAGGCGGTCT GCGCTACGAC CTGCTCAAAC AGGATGCATT
	3951	CGCCGAAAAA GGCAGTGCTT TGGGCTGGAG CGGCAACAGC CTCAGTGAAG
30	4001	GCACGCTGGT CGGACTCGCG GGTCTGAAGC TGTCGCAACC CTTGAGCGAT
	4051	AAAGCCGTC TGTTGCAAC GGCGGGCGTG GAACCGGACC TGAACGGACG
	4101	CGACTACACG GTAACGGCGC GCTTACCGG CGCGACTGCA GCAACCGGCA
	4151	AGACGGGGGC ACGCAATATG CGGCACACCC GTCTGGTTGC CGGCCTGGGC
	4201	GC GGATGTCA AATTGGCAA CGGCTGGAAC GGCTTGGCAC GTTACAGCTA
	4251	CGCCGGTTCC AAACAGTACG GCAACCACAG CGGACGGAGTC GGCCTAGGCT
35	4301	ACCGGTTCT CGAGCACCAC CACCACCAACTGA

40

45

50

55

**EP 1 790 660 A2**

1 MATNDDDVKK AATVAAIAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
51 AADVEADDK GLGLKKVVTN LTAKTVNENKQ NVDAKVAAE SEIEKLTTKL  
101 ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA  
151 DTVDKHAEAF NDIADSDET NTKADEAVKT ANEAKQTAEE TKQNVDAKVK  
201 AAETAAGKAE AAAGTANTAA DKAEEAVAALK TDIKADIATN KDNIAKKANS  
251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEEKSIADH DTRLNGLDKT  
301 VSDLRKETRQ GLAEQAALSG LFQPYNVGRF NVTAAVGGYK SESAVAIGTG  
351 FRFTENFAAK AGVAVGTSSG SSAAYHVGVN YEWGSGGGGT SAPDFNAGGT  
401 GIGSNSRATT AKSAAVSYAG IKNEMCKDRS MLCAGRDDVA VTDRDAKINA  
451 PPPNLHTGDF PPNPDAYKLN INLKPAIEAG YTGRGVEVGI VDTGESVGSI  
501 SFPELYGRKE HGYNENYKNY TAYMRKEAPE DGGGDIEAS FDDEAVIETE  
551 AKPTDIRHK EIGHIDLVSIIIGGRSVDGR PAGGIAPDAT LHIMNTNDT  
601 KNEMMVAIR NAWVKLGERG VRIVNNNSFGT TSRAGTADLF QIANSEEQYR  
651 QALLDYSGGD KTDEGIRLMQ QSDYGNLSYH IRKNKMLFIF STGNDQAQP  
701 NTYALLPFYE KDAQKGIITV AGVDRSGEKF KREMYGEPGT EPLEYGSNHC  
751 GITAMWCLSA PYEASVRFTR TNPIQIAGTS FSAPIVTGTA ALLLQKYPWM  
801 SNDNLRTLL TTAQDIGAVG VDSKFGWGLL DAGKAMNGPA SFPFGDFTAD  
851 TKGTSIAYS FRNDISGTGG LIKKGGSQLQ LHGNNTYTGK TIEEGGLVL  
901 YGNNKSDMRV ETKGALIYNG AASGGSLNSD GIVYLADTDQ SGANETVHIK  
951 GSLQLDGKGT LYTRLGKLLK VDGTAIIGGK LYMSARGKGA GYLNSTGRRV  
1001 PFLSAAKIGQ DYSFTNINET DGGLLASLDS VEKTAGSEGD TLSYYVRRGN  
1051 AARTASAAAH SAPAGLKHAV EQGGSNLENL MVELDASESS ATPETVETAA  
1101 ADRTDMMPGIR PYGATFRAAA AVQHANAADG VRIFNSLAAT VYADSTAABA  
1151 DMQGRRLKAV SDGLDHNGTG LRVIAQTQQD GGTWEQGGVE GKMRGSTQTV  
1201 GIAAKTGENT TAAATLGMGR STWSENSANA KTDISISLFAG IRHDAGDIGY  
1251 LKGLFSYGRY KNSISRSTGA DEHAEGSVNG TLMQLGALGG VNVPFAATGD  
1301 LTVEGGLRYD LLKQDFAEK GSALGWSGNS LTEGTLVGLA GLKLSQPLSD

25

1351 KAVLFATAGV ERDLNGRDYT VTGGFTGATA ATGKTGARNM PHTRLVAGLG  
1401 ADVEFGNGWN GLARYSYAGS KQYGNHSGRV GVGYRFLEHH HHHH\*

30

35

40

45

50

55

961c-ORF46.1

5	1	ATGGCCACAA	ACGACGACGA	TGTTAAAAAA	GCTGCCACTG	TGGCCATTGC
	51	TGCTGCCTAC	AACAATGGCC	AAGAAATCAA	CGGTTCAAA	GCTGGAGAGA
	101	CCATCTACGA	CATTGATGAA	GACGGCACAA	TTACCAAAAA	AGACGCAACT
	151	GCAGCCGATG	TTGAAGCCGA	CGACTTTAAA	GGTCTGGTC	TGAAAAAAAGT
	201	CGTACTAAC	CTGACCAAAA	CCGTCAATGA	AAACAAACAA	AACGTCGATG
	251	CCAAAGTAAA	AGCTGCAGAA	TCTGAAATAG	AAAAGTTAAC	AACCAAGTTA
	301	GCAGACACTG	ATGCCGCTT	AGCAGATACT	GATGCCGCTC	TGGATGCAAC
	351	CACCAACGCC	TTGAATAAT	TGGGAGAAAA	TATAACGACA	TTTGCTGAAG
10	401	AGACTAAGAC	AAATATCGTA	AAAATTGATG	AAAAATTAGA	AGCCGTGGCT
	451	GATACCGTCG	ACAAGCATGC	CGAAGCATTTC	AACGATATCG	CCGATTCTATT
	501	GGATGAAACC	AACACTAAGG	CAGACGAAGC	CGTCAAAACC	GCCAATGAAG
	551	CCAAACAGAC	GGCCGAAGAA	ACCAAAACAAA	ACGTCGATGC	CAAAGTAAAAA
	601	GCTGCAGAAA	CTGCAAGCAGG	CAAAGCCGAA	GCTGCCGCTG	GCACAGCTAA
15	651	TACTGCAGCC	GACAAGGCCG	AAGCTGTCG	TGCAAAAGTT	ACCGACATCA
	701	AAGCTGATAT	CGCTACGAAC	AAAGATAATA	TTGCTAAAAAA	AGCAAAACAGT
	751	GCCGACGTGT	ACACCAAGAGA	AGAGTCTGAC	AGCAAATTG	TCAGAATTGA
	801	TGGTCTGAAC	GCTACTACCG	AAAATTGGA	CACACGCTTG	GCTTCTGCTG
	851	AAAATCCAT	TGCCGATCAC	GATACTCGCC	TGAACGGTTT	GGATAAAACA
	901	GTGTCAGACC	TGCGCAAAGA	AACCCGCCAA	GGCCTTGCAG	AACAAGCCGC
20	951	GCTCTCCGGT	CTGTTCCAAC	CTTACAACGT	GGGTGGATCC	GGAGGAGGAG
	1001	GATCAGATTT	GGCAAACGAT	TCTTTTATCC	GGCAGGTTCT	CGACCGTCAG
	1051	CATTTCGAAC	CCGACGGGAA	ATACCAACCTA	TTCGGCAGCA	GGGGGGAACT
	1101	TGCCGAGCGC	AGCGGCCATA	TCGGATTGGG	AAAATACAA	AGCCATCAGT
	1151	TGGGCAACCT	GATGATTCAA	CAGGCGGCCA	TTAAAGGAAA	TATCGGCTAC
	1201	ATTGTCCGCT	TTTCCGATC	CGGGCACGAA	GTCCATTCCC	CCTTCGACAA
25	1251	CCATGCCTCA	CATTCCGATT	CTGATGAAGC	CGGTAGTCCC	GTGACGGAT
	1301	TTAGCCTTA	CCGCATCCAT	TGGGACGGAT	ACGAACACCA	TCCCGCCGAC
	1351	GGCTATGACG	GGCCACAGGG	CGGGCGCTAT	CCCGCTCCCA	AAGGCGCGAG
	1401	GGATATATAC	AGCTACGACA	AAAAGGGCGT	TGCCCAAAAT	ATCCGCCTCA
	1451	ACCTGACCGA	CAACCGCAGC	ACCGGACAAAC	GGCTTGCCGA	CCGTTTCCAC
30	1501	AATGCCGGA	GTATGCTGAC	GCAAGGAGTA	GGGACGGAT	TCAAACGCGC
	1551	CACCCGATAC	AGCCCCGAGC	TGGACAGATC	GGGCAATGCC	GCCGAAGCCT
	1601	TCAACGGCAC	TGCAAGATATC	GTAAAAAACAA	TCATCGGCGC	GGCAGGAGAA
	1651	ATTGTCGGCG	CAGGCATGTC	CGTGCAGGGC	ATAAGCGAAG	GCTCAAACAT
	1701	TGCTGTCATG	CACGGCTTGG	GTCTGCTTTC	CACCGAAAAC	AAGATGGCGC
	1751	GCATCAACGA	TTTGGCAGAT	ATGGCGCAAC	TCAAAGACTA	TGCCGCAGCA
35	1801	GCCATCCCG	ATTGGGCAGT	CCAAAACCCC	AATGCCGCAC	AAGGCATAGA
	1851	AGCCGTCA	AATATCTTTA	TGGCAGCCAT	CCCCATCAAA	GGGATTGGAG
	1901	CTGTTCGGGG	AAAATACGGC	TTGGGCGGCA	TCACGGCACA	TCCTATCAAG
	1951	CGGTCGCAGA	TGGGCGCGAT	CGCATTGCCG	AAAGGGAAAT	CCGCCGTCA
	2001	CGACAATTTC	GCCGATGCGG	CATA CGCAA	ATACCCGTCC	CCTTACCATT
	2051	CCCGAAATAT	CCGTTCAAAC	TTGGAGCAGC	GTTACGGCAA	AGAAAACATC
40	2101	ACCTCCTCAA	CCGTGCCG	GTCACACGGC	AAAAATGTCA	AACTGGCAGA
	2151	CCAACGCCAC	CCGAAGACAG	GGTACCGTT	TGACGGTAAA	GGGTTCCGA
	2201	ATTTGAGAA	GCACGTGAAA	TATGATACGC	TCGAGCACCA	CCACCACAC
	2251	CACTGA				

45

50

55

1 MATNDDDVKK AATVAIAAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
 51 AADVEADDK GLGLKKVVTN LTKTVNENKQ NVDAKVKAEE SEIEKLTTKL  
 101 ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA  
 151 DTVDKHAEAF NDIADSDET NTKADEAVKT ANEAKQTAEE TKQNVDAKVK  
 201 AAETAAGKAE AAAGTANTAA DKAEEAVAALK TDIKADIATN KDNIAKKANS  
 251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT  
 301 VSDLRKETRQ GLAEQAALSG LFQPYNVGGS GGGGSDLAND SFIRQVLDRQ  
 351 HFEPDGKYHL FGSRGELAER SGHIGLGKIQ SHQLGNLMIQ QAAIKGNIGY  
 401 IVRFSDHGHE VHSPFDNHAS HSDSDEAGSP VDGFSLYRIH WDGYEHHPAD  
 451 GYDGPQGGGY PAPKGARDIY SYDIKGVAQN IRLNLTDRS TGQRLLADRFH  
 501 NAGSMLTQGV GDGFKRATRY SPELDRSGNA AEAFNGTADI VKNIIGAAGE  
 551 IVGAGDAVQG ISEGNSNIAVM HGLGLLSTEN KMARINDLAD MAQLKDYAAA  
 601 AIRDWAVQNP NAAQGIEAVS NIFMAAIPIK GIGAVRGKYG LGGITAHPIK  
 651 RSQMGAIALP KGKSAVSDNF ADAAYAKYPS PYHSRNIRSN LEQRYGKENI  
 701 TSSTVPPSNG KNVKLADQRH PKTGVPDFGK GFPNFEKHVK YDTLEHHHHH  
 15

751 H\*

20

961c-741

1 ATGGCCACAA ACGACGACGA TGTTAAAAAA GCTGCCACTG TGGCCATTGC  
 51 TGCTGCCTAC AACATGGCC AAGAAATCAA CGGTTCAAA GCTGGAGAGA  
 101 CCATCTACGA CATTGATGAA GACGGCACAA TTACCAAAAA AGACGCAACT  
 151 GCAGCCGATG TTGAAGCCGA CGACTTTAAA GGTCTGGTC TGAAAAAAAGT  
 201 CGTGACTAAC CTGACCAAAA CCGTCAATGA AAACAAACAA AACGTCGATG  
 251 CCAAAGTAAA AGCTGCAGAA TCTGAAATAG AAAAGTTAAC ACCAAGTTA  
 301 GCAGACACTG ATGCCGCTT AGCAGATACT GATGCCGCTC TGGATGCAAC  
 351 CACCAACGCC TTGAATAAT TGGGAGAAAA TATAACGACA TTTGCTGAAG  
 401 AGACTAAGAC AAATATCGTA AAAATTGATG AAAAATTAGA AGCCGTGGCT  
 451 GATACCCTCG ACAAGCATGC CGAAGCATTC AACGATATCG CCGATTCTT  
 501 GGATGAAACC AACACTAAGG CAGACGAAGC CGTCAAAACC GCCAATGAAG  
 551 CCAAACAGAC GGCGGAAGAA ACCAAACAAA ACCTGCATGC CAAAGTAAAAA  
 601 GCTCAGAAA CTGCAGCAGG CAAAGCCGAA GCTCCGCTG GCACAGCTAA  
 651 TACTGCAGCC GACAAGGCCG AAGCTGTGCG TGCAAAAGTT ACCGACATCA  
 701 AAGCTGATAT CGCTACGAAC AAAGATAATA TTGCTAAAAA AGCAAAACAGT  
 751 GCCGACGTGT ACACCAGAGA AGAGTCTGAC AGCAAATTG TCAGAATTGA  
 801 TGGTCTGAAC GCTACTACCG AAAAATTGGA CACACGCTTG GCTTCTGCTG  
 851 AAAAATCCAT TGCCGATCAC GATACTCGCC TGAACGGTTT GGATAAAACAA  
 901 GTGTCAGACC TGCGCAAAGA AACCCGCCAA GGCCTTGCAG AACAAGCCGC  
 951 GCTCTCCGGT CTGTTCCAAC CTTACAACGT GGGTGGATCC GGAGGGGGTG  
 1001 GTGTCGCGC CGACATCGGT CGGGGGCTTG CCGATGCACT AACCGCACCG  
 1051 CTCGACCAT AAGACAAAGG TTTGCAGTCT TTGACGCTGG ATCAGTCCGT  
 1101 CAGGAAAAAC GAGAAACTGA AGCTGGCGGC ACAAGGTGCG GAAAAAAACTT  
 1151 ATGGAAACGG TGACAGCCTC AATACGGGCA AATTGAAGAA CGACAAGGTC  
 1201 AGCCGTTTCG ACTTTATCCG CCAAATCGAA GTGGACGGGC AGCTCATTAC  
 1251 CTTGGAGAGT GGAGAGTTCC AAGTATAACAA ACAAAAGCCAT TCCGCCTTAA  
 1301 CCGCCTTCA GACCGAGCAA ATACAAGATT CGGAGCATTC CGGGAAAGATG  
 1351 GTTGCAAAC GCCAGTTCAAG AATCGGCAC ATAGCGGGCG AACATACATC  
 1401 TTTTGACAAG CTTCCCGAAG GCGGCAGGGC GACATATCGC GGGACGGCGT  
 1451 TCGGTTCAGA CGATGCCGGC GGAAACTGA CCTACACCAT AGATTCGCC  
 1501 GCCAAGCAGG GAAACGGCAA AATCGAACAT TTGAAATCGC CAGAACTCAA  
 1551 TGTCGACCTG GCCGCCGCCG ATATCAAGCC GGATGGAAAA CGCCATGCCG  
 1601 TCATCAGCGG TTCCGTCCTT TACAACCAAG CCGAGAAAGG CAGTTACTCC  
 1651 CTCGGTATCT TTGGCGGGAAA AGCCCAGGAA GTTGCAGGCA GCGCGGAAGT  
 1701 GAAAACCGTA AACGGCATAAC GCCATATCGG CCTTGCCGCC AAGCAACTCG  
 1751 AGCACCACCA CCACCAACCA TGA

55

1 MATNDDDVKK AATVAIAAAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
 51 AADVEADDK GLGLKKVVTN LTAKTVENKQ NVDAVKVAAE SEIEKLTTKL  
 101 ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA  
 151 DTVDKHAEAF NDIADSLDET NTKADEAVKT ANEAQTAEE TKQNVDAKVK  
 201 AAETAAGKAE AAAGTANTAA DKAEAVAACK TDIKADIATN KDNIAKKANS  
 251 ADVYTREESD SKFVRIDGLN ATTEKLDTRL ASAEKSIADH DTRLNGLDKT  
 301 VSDLRKETRQ GLAEQAALSG LFQPYNVGGS GGGGVAADIG AGLADALTAP  
 351 LDHKDKGLQS LTLDQSVRKN EKLKLAQGA EKTYGNGDSL NTGKLKNDKV  
 401 SRFDFIROIE VDGQLITLIES GEFQVYKOSH SALTAFQTEQ IODSEHSGKM  
 451 VAKRQFRIGD IAGEHTSFDR LPEGGRATYR GTAFCGSDAG GKLTYTIDFA  
 501 AKQGNGKIEH LKSPELVNDL AAADIKPDGK RHAVISGSVL YNQAEKGYS  
 551 LGIFGGKAQE VAGSAEVKTV NGIRHIGLAA KQLEHHHHH \*

15

961c-983

1 ATGGCCACAA ACGACGACGA TGTTAAAAAA GCTGCCACTG TGGCCATTGC  
 51 TGCTGCCTAC AACAAATGCC AAGAAATCAA CGGTTCAAA GCTGGAGAGA  
 101 CCATCTACGA CATTGATGAA GACGGCACAA TTACCAAAAA AGACGCAACT  
 151 GCAGCCGATG TTGAAGCCGA CGACTTTAAA GGTCTGGTC TGAAAAAAAGT  
 201 CGTACTAAC CTGACCAAAA CCGTCAATGA AAACAAACAA AACGTCGATG  
 251 CCAAAGTAAA AGCTGCAGAA TCTGAAATAG AAAAGTTAAC ACCAAGTTA  
 301 GCAGACACTG ATGCCGCTTT AGCAGATACT GATGCCGCTC TGGATGCAAC  
 351 CACCAACGCC TTGAATAAAAT TGGGAGAAAA TATAACGACA TTGCTGAAG  
 401 AGACTAAGAC AAATATCGTA AAAATTGATG AAAAATTAGA AGCCGTGGCT  
 451 GATACCGTCG ACAAGCATGC CGAACGATTC AACGATATCG CCGATTCAATT  
 501 GGATGAAACC AACACTAAGG CAGACGAAGC CGTCAAAACC GCCAATGAAG

30

35

40

45

50

55

5	551	CCAAACAGAC GGCGAAGAA ACCAACAAA ACGTCGATGC CAAAGTAAAA 601 GCTGCAGAAA CTGCAGCAGG CAAAGCCGAA GCTGCCGCTG GCACAGCTAA 651 TACTGCAGCC GACAAGGCGG AAGCTGTCCG TGCAAAAGTT ACCGACATCA 701 AAGCTGATAT CGCTACGAAC AAAGATAAA TTGCTAAAAA AGCAAACAGT 751 GCCGACGTGT ACACCAAGA AGAGTCGAC AGCAAATTG TCAGAATTGA 801 TGGTCTGAAC GCTACTACCG AAAAATTGCA CACACGCTTG GCTTCTGCTG 851 AAAAATCCAT TGCCGATCAC GATACTCGCC TGAAACGGTT GGATAAAACA 901 GTGTCAAGACC TGCGCAAAGA AACCCGCCAA GGCCTTGAG AACAAAGCCGC 951 GCTCTCCGGT CTGTTCCAAC CTTACAACGT GGGTGGATCC GGCGGAGGCG 1001 GCACTCTGTC GCCCGACTTC AATGCAGGGC GTACCGGTAT CGGCAGCAAC 1051 AGCAGACCAA CAACAGCGAA ATCAGCAGCA GTATCTTACG CCGGTATCAA 1101 GAACGAAATG TGCAAAGACA GAAGCATGCT CTGTGCCGGT CGGGATGACG 1151 TTGCGGTTAC AGACAGGGAT GCCAAATCA ATGCCCCCCC CCCGAATCTG 1201 CATACCCGAG ACTTTCCAAA CCCAAATGAC GCATACAAAGA ATTGATCAA 1251 CCTCAAACCT GCAATTGAAG CAGGCTATAC AGGACGCCGG GTAGAGGTAG 1301 GTATCGTCGA CACAGGCGAA TCCGTCGCGA GCATATCCTT TCCCGAACTG 1351 TATGCCAGAA AAGAACACGG CTATAACGAA ATTACAAAAA ACTATACGGC 1401 GTATATGCGG AAGGAAGCGC CTGAAGACCG AGGCGGTAAA GACATTGAAG 1451 CTTCTTCGA CGATGAGGCC GTTATAGAGA CTGAAGCAAA GCGGACGGAT 1501 ATCCGCCACG TAAAAGAAAT CGGACACATC GATTGKTCT CCCATATTAT 1551 TGGCGGGCGT TCCGTGGACG GCAGACCTCC AGGCCTGATT GCGCCCGATG 1601 CGACGCTACA CATAATGAAT ACGAATGATG AAACCAAGAA CGAAATGATG 20 1651 GTTGCAGCCA TCCGCAATGC ATGGGTCAAG CTGGGCGAAC GTGGCGTGC 1701 CATCGTCAT AACAGTTTG GAACAACATC GAGGGCAGGC ACTGCCGACC 1751 TTTTCCCAAAT AGCCAATTCC GAGGAGCAGT ACCGCCAAGC GTTGCTCGAC 1801 TATTCCGGCG GTGATAAAAC AGACGAGGGT ATCCGCTGA TGCAACAGAG 1851 CGATTACGGC AACCTGTCT ACCACATCCG TAATAAAAAC ATGTTTTCA 1901 TCTTTTCGAC AGGCAATGAC GCACAAGCTC AGCCCAACAC ATATGCCCTA 25 1951 TTGCCATTTC ATGAAAAAAGA CGCTAAAAAA GGCATTATCA CAGTCGCAGG 2001 CGTAGACCGC AGTGGAGAAA AGTTCAAAACG GGAAATGTAT GGAGAACCGG 2051 GTACAGAACC GCTTGAGTAT GGCTCCAACC ATTGCGGAAT TACTGCCATG 2101 TGGTGCCTGT CGGCACCCCTA TGAAGCAAGC GTCCGTTCA CCCGTACAAA 2151 CCCGATTCAA ATTGCGGGAA CATCCITTC CGCACCCATC GTAACCGGGCA 2201 CGCGGCTCT GCTGCTGCAG AAATACCGT GGATGAGCAA CGACAACCTG 2251 CGTACACCGT TGCTGACGAC GGCTCAGGAC ATCGGTGCAG TCGCGTGG 30 2301 CAGCAAGTTC GGCTGGGGAC TGCTGGATGC GGGTAAGGCC ATGAAACGGAC 2351 CCGCTCCTT TCCGTTCGGC GACTTTACCG CCGATACGAA AGGTACATCC 2401 GATATTGCCT ACTCCTCCG TAACGACATT TCAGGCACGG CGGGCCTGAT 2451 CAAAAAAAGGC GGCGACCAAC TGCAACTGCA CGGCAACAAAC ACCTATACGG 2501 GCAAAACCAT TATCGAAGGC GGTTCGCTGG TGTGTAACGG CAACAACAAA 2551 TCGGATATGC GCGTCGAAAC CAAAGGTGCC CTGATTATA ACGGGGCGCC 2601 ATCCGGCGC AGCCTGAACA GCGACGGCAT TGTCTATCTG GCAGATACCG 2651 ACCAATCCGG CGCAAACGAA ACCGTACACA TCAAAGGCAG TCTGCAGCTG 2701 GACGGCAAAAG CTACGCTGTA CACACCTTGC CGCAAACACTGC TGAAACTGGA 2751 CGGTACGGCG ATTATCGGGC GCAAGCTGTA CATGTCGGCA CGGGCGAAGG 2801 GGGCAGGCTA TCTCAACAGT ACCGGACGAC GTGTTCCCTT CCTGAGTGCC 2851 GCAAAATCCG CGCAGGATTA TTCTTCTTC ACAAAACATCG AAACCCACGG 40 2901 CGGCCTGCTG GCTCCCTCG ACAGCGTCGA AAAAACAGCG GGCAGTGAAG 2951 GCGACACGCT GTCCATTAT GTCCGTCGCG GCAATGCGGC ACGGACTGCT 3001 TCGGCAGCGG CACATTCCGC GCCCAGCGGT CTGAAACACG CCCTACAAACA 3051 GGGCGGCAGC AACTCTGGAA ACCTGATGGT CGAAACTGGAT GCCTCCGAAT 3101 CATCCGCAAC ACCCGAGACG GTTGAACATG CGGCAGCCGA CGGCACAGAT 3151 ATGCCGGCA TCCGCCCCCTA CGCGCAACT TTCCGCGCAG CGGCAGCCGT 45 3201 ACAGCATGCG AATGCCGGCG ACGGTGTACG CATCTTCAAC AGTCTCGCG 3251 CTACCGCTA TGCCGACAGT ACCGCGCCCG ATGCCGATAT GCAGGGACGC 3301 CGCCTGAAAG CGGTATCGGA CGGGTTGGAC CACAACGGCA CGGGTCTGCG 3351 CGTCATCGCG CAAACCCAAAC AGGACGGTGG AACGTGGAA CAGGGCGGTG 3401 TTGAAGGCAA AATGCCGGC AGTACCCAAA CGCTCGGCAT TGCCCGAAGA 3451 ACCGGCAGAA ATACGACAGC AGCCGCCACA CTGGGCATGG GACCGCAGCAC 50 3501 ATGGACCGAA AACAGTGCAA ATGCAAAAC CGACAGCATT AGTCTGTTTG 3551 CAGGCATACG GCACGATGCG GCGGATATCG GCTATCTCAA AGGCCTGTT 3601 TCCTACGGAC GCTACAAAAA CAGCATCAGC CGCAGCACCG GTGCGGACGA 3651 ACATGCCGA GGCAGCGTCA ACGGCACGCT GATGCAGCTG GGCGCACTGG 3701 GCGGTGCTAA CGTCCCGTT GCGCAGACGG GAGATTGAC GGTGCAAGGC 3751 GGTCTCGCGT ACGACCTGCT CAAACAGGAT GCATTGCCG AAAAAGGCAG 3801 TGCTTGGGC TGGAGCGGGCA ACAGCCTCAC TGAAGGCACG CTGGTGGAC 55 3851 TCGCGGGTCT GAAAGCTGTCG CAACCCCTGA GCGATAAAGC CGTCCTGTTT
---	-----	---

5           3901 GCAACGGCGG GCGTGGAACG CGACCTGAAC GGACCGCACT ACACGGTAAC  
      3951 GGGCGGCTTT ACCGGCGCGA CTGCAGCAAC CGGCAAGACG GGGGCACGCA  
      4001 ATATGCCGCA CACCCGTCTG GTTGCCGGCC TGGGCGCGGA TGTCGAATT  
      4051 GGCAACGGCT GGAACGGCTT GGCACTGTTAC AGCTACGCCG GTTCCAAACA  
      4101 GTACGGCAAC CACAGCGGAC GAGTCGGCGT AGGCTACCAGG TTCCCTCGAGC  
      4151 ACCACCACCA CCACCAACTGA

10           1 MATNDDDVKK AATVAIAAY NNGQEINGFK AGETIYDIDE DGTITKKDAT  
      51 AADVEADDK GLGLKKVVTN LTKTVNENKQ NVDAVKAAE SEIEKLTTKL  
      101 ADTDAALADT DAALDATTNA LNKLGENITT FAEETKTNIV KIDEKLEAVA  
      151 DTVDKHAEAF NDIAIDLDET NTKADEAVKT ANEAKQTAEE TKQNVDAKVK  
      201 AAETAAGKAE AAAGTANTAA DKAEAVAAKV TDIKADIATN KDNIAKKANS  
      251 ADVYTREESD SKFVRIDGLN ATTEKLDRRL ASAEKSIADH DTRLNGLDKT  
      301 VSDLRKETRQ GLAEQAALSG LFQPYNVGGS GGGGTSAPDF NAGGTGIGSN  
      351 SRATTAKSAA VSYAGIKNEM CKDRSMLCAG RDDVAVTDRD AKINAPPNL  
      401 HTGDFPNPND AYKNLINLKP AIEAGYTGRG VEVGIVDTGE SVGSISFPEL  
      451 YGRKEHGYNE NYKNYTAYMR KEAPEDGGK DIEASFDEA VIETEAKPTD  
      501 IRHVKEIGHI DLVSHIIGGR SVDGRPAGGI APDATLHIMN TNDETKNEMM  
      551 VAAIRNAWK LGERGVRIVN NSFGTTSRAG TADLFQIANS EEQYRQALLD  
      601 YSGGDKTDEG IRLMQQSDYG NLSYHIRNKN MLFIFSTGND AQAQPNTYAL  
      651 LPFYEKDAQK GIITVAGVDR SGEKFKREMY GEPGTEPLEY GSNHCGITAM  
      701 WCLSAPYEAS VRFTRTNPIQ IAGTSFSAPI VTGTAALLQ KYPWMSNDNL  
      751 RTTLLTTAQD IGAVGVDSKF GWGLLDAGKA MNGPASFPFG DFTADTKGTS  
      801 DIAYSFRNDI SGTGGLIKKG GSQQLHLGNN TYTGKTIIEG GSLVLYGNNK  
      851 SDMRVETKGA LIYNGAASGG SLNSDGIVYL ADTDQSGANE TVHIKGSLQL  
      901 DGKGTLYTRL GKLLKVDGTA IIIGGKLYMSA RGKGAGYLN TGRRVPFLSA  
      951 AKIGQDYSFF TNIETDGGLL ASLDSVEKTA GSEGDTLSYY VRGNAAARTA  
      1001 SAAAHSAPAG LKHAVEQGGS NLLENLMVELD ASESSATPET VETAAADRTD  
      1051 MPGIRPYGAT FRAAAAVQHA NAADGVRIFN SLAATVYADS TAAHADMQGR  
      1101 RLKAVSDGLD HNGTGLRVIA QTQQDGGGTWE QGGVEGKMRC STQTVGIAAK  
      1151 TGENTTAAAT LGMGRSTWSE NSANAKTDI SLFAGIRHDA GDIGYLKGLF  
      1201 SYGRYKNSIS RSTGADEHAE GSVNGTLMQL GALGGVNVPF AATGDLTVEG  
      1251 GLRYDLLKQD AFAEKGSALG WSGNSLTEGT LVGLAGLKLS QPLSDKAVLF  
      1301 ATAGVERDLN GRDYTVTGGF TGATAATGKT GARNMPHTRL VAGLGADVEF  
      1351 GNGWNGLARY SYAGSKQYGN HSGRVGVGYR FLEHHHHHH\*

35

40

45

50

55

961cL-ORF46.1

5            1 ATGAAACACT TTCCATCCAA AGTACTGACC ACAGCCATCC TTGCCACTTT  
       51 CTGTAGCGGC GCACTGGCAG CCACAAACGA CGACGATGTT AAAAAAAGCTG  
       101 CCACTGTGGC CATTGCTGCT GCCTACAACA ATGGCCAAGA AATCAACGGT  
       151 TTCAAAAGCTG GAGAGACCAT CTACGACATT GATGAAGACG GCACAATTAC  
       201 CAAAAAAAGAC GCAACTGCAG CCGATGTTGA AGCCGACGAC TTTAAAGGTC  
       251 TGGGTCTGAA AAAAGTCGTG ACTAACCTGA CCAAAACCGT CAATGAAAAC  
       301 AAACAAAACG TCGATGCCAA AGTAAAAGCT GCAGAATCTG AAATAGAAAA  
       351 GTTAACAACCG AAGTTAGCAG ACACTGATGC CGCTTAGCA GATACTGATG  
       401 CCGCTCTGGA TGCAACCACC AACGCCTTGA ATAAATTGGG AGAAAATATA  
       451 ACGACATTTG CTGAAGAGAC TAAGACAAAT ATCGTAAAAA TTGATGAAAA  
       501 ATTAGAACGC GTGGCTGATA CCGTCGACAA GCATGCCGAA GCATTCAACG  
       551 ATATCGCCGA TTCATTGGAT GAAACCAACA CTAAGGCAGA CGAAGCCGTC  
       601 AAAACCGCCA ATGAAGCCAA ACAGACGGCC GAAGAAACCA AACAAAACGT  
       651 CGATGCCAAA GTAAAAGCTG CAGAAACTGTC AGCAGGCAA GCCGAAGCTG  
       701 CCGCTGGCAC AGCTAATACT GCAGCCGACA AGGCCGAAGC TGTCGCTGCA  
       751 AAAGTTACCG ACATCAAAGC TGATATCGCT ACGAACAAAG ATAATATTGC  
       801 TAAAAAAAGCA AACAGTGGCG ACGTGTACAC CAGAGAAGAG TCTGACAGCA  
       851 AATTGTCAG AATTGATGGT CTGAACGCTA CTACCGAAAA ATTGGACACA  
       901 CGCTTGGCTT CTGCTGAAAA ATCCATTGCC GATCACGATA CTCGCCTGAA  
       951 CGGTTTGGAT AAAACAGTGT CAGACCTGCG CAAAGAAACC CGCCAAGGCC  
 20            1001 TTGCAGAACAA AGCCCGCGTC TCCGGTCTGT TCCAACCTTA CAACGTGGGT  
       1051 GGATCCGGAG GAGGAGGATC AGATTTGGCA AACGATTCTT TTATCCGGCA  
       1101 GGTTCTCGAC CGTCAGCATT TCGAACCCGA CGGGAAATAC CACCTATTG  
       1151 GCAGCAGGGG GGAACATTGCC GAGCGCAGCG GCCATATCGG ATTGGGAAAA  
       1201 ATACAAAGCC ATCAGTTGGG CAACCTGATG ATTCAACAGG CGGCCATTAA  
       1251 AGGAAATATC GGCTCACATTG TCCGCTTTTC CGATCACGGG CACGAAGTCC  
       1301 ATTCCCCCTT CGACAAACCAT GCCTCACATT CCGATTCTGA TGAAGCCGGT  
       1351 AGTCCCCTTG ACGGATTTAG CCTTTACCGC ATCCATTGGG ACGGATAACGA  
       1401 ACACCATCCCC GCCGACGGCT ATGACGGGCC ACAGGGCGGC GGCTATCCCG

30

35            1451 CTCCCAAAGG CGCGAGGGAT ATATACAGCT ACGACATAAA AGGC GTTGCC  
       1501 CAAAAATATCC GCCTCAACCT GACCGACAAC CGCAGCACCG GACAACGGCT  
       1551 TGCCGACCGT TTCCACAATG CCGGTAGTAT GCTGACGCAA GGAGTAGGCG  
       1601 ACGGATTCAA AC CGCC ACC CGATA CAGCC CCGAGCTGG CAGATCGGGC  
       1651 AATGCCGCCG AAGCCTCAA CGGC ACTGCA GATATCGTT AAAACATCAT  
       1701 CGGCGCGGC GGAGAAATTG TCGGCGCAGG CGATGCCGT CAGGGCATAA  
       1751 GCGAAGGCTC AAA CATTGCT GTCATGCACG GCTTGGGTCT GCTTTCCACC  
       1801 GAAAACAAGA TGGCGCGCAT CAACGATTG GCAGATATGG CGCAACTCAA  
       1851 AGACTATGCC GCAGCAGCCA TCCGCGATG GGCAGTCCAA AACCCCAATG  
       1901 CCCGACAAAGG CATAGAACCC GTCAGCAATA TCTTTATGGC AGCCATCCCC  
       1951 ATCAAAGGG A TTGGAGCTGT TCAGGGAAAAA TACGGCTTGG GCGGCATCAC  
       2001 GGCACATCCT ATCAAGCGGT CGCAGATGGG CGCGATCGCA TTGCCGAAAG  
       2051 GGAAATCCGC CGTCAGCGAC AATTTGCCG ATGCGGCATA CGCCAAATAC  
       2101 CCGTCCCCCTT ACCATTCCCG AAATATCCGT TCAAACCTGG AGCAGCGTTA  
       2151 CGGCAAAGAA AACATCACCT CCTCAACCGT GCCGCCGTCA AACGGAAAAA  
       2201 ATGTCAAACCT GGCAGACCAA CGCCACCCGA AGACAGGCCT ACCGTTTGAC  
       2251 GGTAAAGGGT TTCCGAATT TGAGAACAC GTGAAATATG ATACGTAACCT  
       2301 CGAG

50

55

1 MKHFPSKVLT TAILATFCSG ALAATNDDDV KKAATV рівні AYNNGQEING  
 51 FKAGETIYDI DEDGTITKKD ATAADVEADD FKGLGLKKVV TNLTKTVNEN  
 101 KQNVDAKVKA AESEIEKLTT KLADTDAAALA DTDAALDATT NALNKLGENI  
 151 TTFAEETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLL ETNTKADEAV  
 201 KTANEAKQTA EETKQNVDAK VKAAETAAGK AEEAAAGTANT AADKAEAVAA  
 251 KVTDIKADIA TNKDNIAKKA NSADVYTREE SDSKFVRIDG LNATTEKLDT  
 301 RLASAEKSIA DHDRNLNGLD KTVSDLRKET RQGLAEQAAL SGLFQPYNVG  
 351 GSGGGSDLA NDSFIRQVLD RQHFEPDGKY HLFGRGELA ERSGHIGLGK  
 401 IQSHQLGNLM IQQAAIKGNI GYIVRFSDHG HEVHSFVDNH ASHSDSDEAG  
 451 SPVDFSLYR IHWDGYEHHP ADGYDGPQGG GYPARKGARD IYSYDIKGVA  
 501 QNIRLNITDN RSTGQRLLADR FHNAGSMLTQ GVGDGFKRAT RYSPELDRSG  
 551 NAAEAFNGTA DIVKNIIGAA GEIVGAGDAV QGISEGSNIA VMHGLGLLST  
 601 ENKMARINDL ADMAQLKDYA AAAIRDWAVQ NPNAAQGIEA VSNIFMAAIP  
 651 IKGIGAVRGK YGLGGITAHP IKRSQMGAIА LPKGKSAVSD NFADAAYAKY  
 701 PSPYHSRNIR SNLEQRYGKE NITSSTVPPS NGKNVKLADQ RHPKTGVPDF  
 751 GKGFPNFEKH VKYDT\*

**961cL-741**

20 1 ATGAAACACT TTCCATCCAA AGTACTGACC ACAGGCCATCC TTGCCACTTT  
 51 CTGTAGCGGC GCACTGGCAG CCACAAACGA CGACGATGTT AAAAAAGCTG  
 101 CCACTGTGGC CATTCGCTGC GCCTACACA ATGGCCAAGA AATCAACGGT  
 151 TTCAAAGCTG GAGAGACCAT CTACGACATT GATGAAGACG GCACAATTAC  
 201 CAAAAAAGAC GCAACTGCAG CCGATGTTGA AGCCGACGAC TTTAAAGGTC  
 251 TGGGCTGAA AAAAGTCGTG ACTAACCTGA CCAAAACCGT CAATGAAAAC  
 301 AAACAAAACG TCGATGCCA AGTAAAAGCT GCAGAATCTG AAATAGAAAA  
 351 GTTAACAACC AAGTTAGCAG AACTGATGC CGCTTAGCA GATACTGATG  
 401 CCGCTCTGGA TGCAACCACC AACGCCCTGA ATAAATTGGG AGAAAATATA  
 451 ACGACATTTG CTGAAGAGAC TAAGACAAAT ATCGTAAAAAA TTGATGAAAA  
 501 ATTAGAAGCC GTGGCTGATA CCGTCGACAA GCATGCCGAA GCATTCAACG  
 551 ATATCGCCGA TTCATTGGAT GAAACCAACA CTAAGGCAGA CGAACGCGTC  
 601 AAAACCGCCA ATGAAGCCAA ACAGACGGCC GAAGAAACCA AACAAAACGT  
 651 CGATGCCAA GTAAAAGCTG CAGAAACTGC AGCAGGCAA GCCGAAGCTG  
 701 CCGCTGGCAC AGCTAATACT GCAGCCGAGC AGGCCGAAGC TGTCCCTGCA  
 751 AAAGTTACCG ACATCAAAGC TGATATCGCT ACAGAACAAAG ATAATATTGC  
 801 TAAAAAAGCA AACAGTGCCG ACGTGTACAC CAGAGAAGAG TCTGACAGCA  
 851 AATTGTCAG AATTGATGGT CTGAACGCTA CTACCGAAAA ATTGGACACA  
 901 CGCTTGGCTT CTGCTGAAAA ATCCATTGCC GATCACGATA CTCGCCTGAA  
 951 CGGTTGGAT AAAACAGTGT CAGACCTGCG CAAAGAAACC CGCCAAGGCC  
 1001 TTGCAGAAC AGCCCGCTC TCCGGTCTGT TCCACCTTA CAACGTGGGT  
 1051 GGATCCGGAG GGGTGGTGT CGCCGCCGAC ATCGGTGCGG GGCTTGCCGA  
 1101 TGCACTAACC GCACCGCTCG ACCATAAAGA CAAAGTTTG CAGTCTTTGA  
 1151 CGCTGGATCA GTCCGTCAAGG AAAAACGAGA AACTGAAGCT GGCAGCACAA  
 1201 GGTGCGAAAA AAACATTATGG AAACGGTGAC AGCCTCAATA CGGGCAAATT  
 1251 GAAGAACGAC AAGGTCAAGCC GTTTCGACTT TATCCGCAA ATCGAACGTGG  
 1301 ACGGGCAGCT CATTACCTTG GAGAGTGGAG AGTCCAAGT ATACAAACAA  
 1351 AGCCATTCCG CCTTAACCGC CTTTCAGACC GAGCAAATAC AAAGATTCCGA  
 1401 GCATTCCGGG AAGATGGTTG CGAAACGCCA GTTCAGAATC GGCAGACATAG

50 1451 CGGGCGAACAA TACATCTTT GACAAGCTTC CCGAAGGCCGG CAGGGCGACA  
 1501 TATCGCGGGA CGCGGTTCGG TTTCAGACGAT GCCGGCGGAA AACTGACCTA  
 1551 CACCATAGAT TTGCGCGCCA AGCAGGGAAA CGGCAAATC GAACATTGAA  
 1601 AATCGCCAGA ACTCAATGTC GACCTGGCCG CGGCCGATAT CAAGCCGGAT  
 1651 GGAAAACGCC ATGCCGTCAAT CAGCGGTTCC GTCCCTTACA ACCAACGCCA  
 1701 GAAAGGCAGT TACTCCCTCG GTATCTTGG CGGAAAAGCC CAGGAAGTTG  
 1751 CGGGCAGCGC GGAAGTGAAA ACCGTAAACG GCATACGCCA TATCGGCCTT  
 1801 GCCGCCAACG AACTCGAGCA CCACCAAC CACCACTGA

5  
1 MKHFPSKVLT TAILATFCSG ALAATNDDV KKAATVAIAA AYNNGQEING  
51 FKAGETIYDI DEDGTITKKD ATAADVEADD FKGLGLKKVV TNLTKTVNEN  
101 KQNVDAVKVA AESEIEKLTT KLADTDAAALA DTDAALDATT NALNKLGENI  
151 TTFAEETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLL ETNTKADEAV  
201 KTANEAKQTA EETKQNVDAK VKAAETAAGK AEEAAAGTANT AADKAEAVAA  
251 KVTDIKADIA TNKDNIACKA NSADVYTREE SDSKVF RIDG LNATTEKLDT  
301 RLASAEKSIA DHDTRLNGLD KTVSDLRKET RQGLAEQAAL SGLFQPYNVG  
351 GSGGGVVAAD IGAGLADALT APLDHDKDGL QSLTLQSVR KNEKLKLAAQ  
401 GAEKYTGNGD SLNTGKLND KVSRFDIFRQ IEVDGQLITL ESGEFQVYKQ  
451 SHSALTAFQT EQIQDSEHSG KMVAKRQFR GDIAGEHTSF DKLPEGGRAT  
501 YRGTAFGSDD AGGKLTYTID FAAKQGNKGK EHLKSPELNV DLAAADIKPD  
551 GKRHAVISGS VLYNQAEKGS YSLGIFGGKA QEVAGSAEVK TVNGIRHIGL  
601 AAKQLEHHHH HH\*

15

961cL-983

20  
1 ATGAAAACACT TTCCATCCAA AGTACTGACC ACAGCCATCC TTGCCACTTT  
51 CTGTAGCGGC GCACTGGCAG CCACAAACGA CGACGATGTT AAAAAAGCTG  
101 CCACTGTGGC CATTGCTGCT GCCTACAACA ATGGCCAAGA AATCAACGGT  
151 TTCAAAGCTG GAGAGACCAT CTACGACATT GATGAAGACG GCACAATTAC  
201 CAAAAAAAGAC GCAACTGCAG CCGATGTTGA AGCCGACGAC TTTAAAGGTC  
251 TGGGTCTGAA AAAAGTCGTG ACTAACCTGA CCAAAACCGT CAATGAAAAC  
301 AAACAAAACG TCGATGCCA AGTAAAAGCT GCAGAACTCG AAATAGAAAA  
351 GTTAACAAACC AAGTTAGCAG AACTGTGATGC CGCTTAGCA GATACTGATG  
401 CCGCTCTGGA TGCAACCACC AACGCCTTGA ATAAATTGGG AGAAAATATA  
451 ACGACATTG CTGAAAGAGAC TAAGACAAAT ATCGTAAAAA TTGATGAAAAA  
501 ATTAGAAGCC GTGGCTGATA CGTCGACCAA GCATGCCGAA GCATTCAACG  
551 ATATCGCCGA TTCAATTGGAT GAAACCAACA CTAAGGCAGA CGAAGCCGTC  
601 AAAACGCCA ATGAAGCCAA ACAGACGGCC GAAGAAAACCA AACAAAACGT  
651 CGATGCCAAA GTAAAAGCTG CAGAAACTGC AGCAGGCAAA GCCGAAGCTG  
701 CCGCTGGCAC AGCTAATACT GCAGCCGACA AGGCCGAAGC TGTCGCTGCA  
751 AAAGTTACCG ACATCAAAGC TGATATCGCT ACGAACAAAG ATAATATTGC  
801 TAAAAAAAGCA AACAGTGCCG ACGTGTACAC CAGAGAAGAG TCTGACAGCA  
851 AATTGTCAG AATTGATGGT CTGAACGCTA CTACCGAAAA ATTGGACACA  
901 CGCTTGGCTT CTGCTGAAAA ATCCATTGCC GATCACGATA CTCGCCTGAA  
951 CGGTTGGAT AAAACAGTGT CAGACCTCGC CAAAGAAACC CGCCAAGGCC  
35  
1001 TTGCAGAACAA AGCCCGCTC TCCGGTCTGT TCCACCTTA CAACGTGGGT  
1051 GGATCCGGCG GAGGCAGCAC TTCTGCGCCC GACTTCAATG CAGGCGGTAC  
1101 CGGTATCGGC AGCAACAGCA GAGCAACAAAC AGCGAACATCA GCAGCAGTAT  
1151 CTTACGCCGG TATCAAGAAC GAAATGTGCA AAGACAGAAC CATGCTCTGT  
1201 GCCGGTCGGG ATGACGTTGC GGTTACAGAC AGGGATGCC AAATCAATGC  
1251 CCCCCCCCCCG AATCTGCATA CCGGAGACTT TCCAAACCCA AATGACGCAT  
40  
1301 ACAAGAATTG GATCAACCTC AAACCTGCAA TTGAAGCAGG CTATACAGGA  
1351 CGCGGGGTAG AGGTAGGTAT CGTCGACACA GGCGAATCCG TCGGCAGCAT  
1401 ATCCTTTCCC GAACGTATG GCAGAAAAGA ACACGGCTAT AACGAAAATT  
1451 ACAAAAACCA TACGGCGTAT ATGCGGAAGG AAGGCCCTGA AGACGGAGGC  
1501 GGTAAAGACA TTGAAGCTTC TTTCGACGAT GAGGGCTTA TAGAGACTGA  
1551 AGCAAGCCG ACGGATATCC GCCACGTAAA AGAAAATCGGA CACATCGATT  
45  
1601 TGGTCTCCCA TATTATTGGC GGGCGTCCG TGGACGGCAG ACCTGCAGGC  
1651 GGTATTGCGC CCGATGCGAC GCTACACATA ATGAATACGA ATGATGAAAC  
1701 CAAGAACGAA ATGATGGTTG CAGCCATCCG CAATGCATGG GTCAAGCTGG  
1751 GCGAACGTGG CGTGCATC GTCAATAACA GTTTGGAAC AACATCGAGG  
1801 GCAGGCACTG CCGACCTTT CCAAATAGCC AATTGGAGG AGCAGTACCG  
1851 CCAAGCGTTG CTCGACTATT CCGGCAGGTGA TAAAACAGAC GAGGGTATCC  
1901 GCCTGATGCA ACAGAGCGAT TACGGCAACC TGTCTACCA CATCCGTAAT  
1951 AAAAACATGC TTTTCATCTT TTGACAGGC AATGACGCAC AAGCTCAGCC  
50  
2001 CAACACATAT GCCCTATTGC CATTATGAA AAAAGACGCT CAAAAAGGCA  
2051 TTATCACAGT CGCAGCGTA GACCGCAGTG GAGAAAAGTT CAAACGGGAA

55

5	2101	ATGTATGGAG AACCGGGTAC AGAACCGCTT GAGTATGGCT CCAACCATTG
	2151	CGGAATTACT GCCATGTGGT GCCTGTCGGC ACCCTATGAA GCAAGCGTCC
	2201	GTTTCACCCG TACAAACCCG ATTCAAATTG CCGGAACATC CTTTCCGCA
	2251	CCCATCGTAA CGGCACGGC GGCTCTGCTG CTGCAGAAAT ACCCGTGGAT
	2301	GAGCAACGAC AACCTCGTAA CCACGTTGCT GACGACGGCT CAGGACATCG
	2351	GTGCAGTCGG CGTGGACAGC AAGTTCGGCT GGGGACTGCT GGATGCGGGT
	2401	AAGGCCATGA ACGGACCCGC GTCCCTTCGG TTCCGGCAGT TTACCGCCGA
	2451	TACGAAAGGT ACATCCGATA TTGCCTACTC CTTCCGTAAC GACATTTCAAG
	2501	GCACGGCGG CCTGATCAAA AAAGGCGGCA GCCAAGTGCAG ACTGCACGGC
10	2551	AACAACACCT ATACGGGCAA AACCAATTATC GAAGGCGGTT CGCTGGTGT
	2601	GTACGGCACAC AACAAATCGG ATATGCGCT CGAAACAAA GGTGCGCTGA
	2651	TTTATAACGG GGCGGCATCC GGCGGCAGCC TGAACAGCGA CGGCATTGTC
	2701	TATCTGGCAG ATACCGACCA ATCCGGCGCA AACGAAACCG TACACATCAA
	2751	AGGCAGTCTG CAGCTGGACG GCAAAGGTAC GCTGTACACA CGTTTGGGCA
	2801	AACTGCTGAA AGTGGACGGT ACGGCGATTA TCGGCGCAA GCTGTACATG
15	2851	TCGGCACCGC GCAAGGGGGC AGGCTATCTC AACAGTACCG GACGACGTGT
	2901	TCCCTTCTG AGTGCAGGCC AAATCGGGCA GGATTATTCT TTCTTCACAA
	2951	ACATCGAAAC CGACGGCGGC CTGCTGGCTT CCCTCGACAG CGTCGAAAAA
	3001	ACAGCGGGCA GTGAAGGCGA CACGCTGTCC TATTATGTCC GTCGCGGCAA
	3051	TGCGGCACGG ACTGCTTCGG CAGCGGCACA TTCCCGGCCG GCCGGTCTGA
20	3101	AACACGGCGT AGAACACAGGC GGCAGCAATC TGGAAAAACCT GATGGTCGAA
	3151	CTGGATGCCT CCGAATCATC CGCAACACCC GAGACGGTTG AAACTGCAGC
	3201	AGCCGACCGC ACAGATATGCC CGGGCATCGG CCCCTACGGC GCAACTTTCC
	3251	GCGCAGCGGC AGCCGTACAG CATCGGAATG CCGCCGACGG TGTACGCATC
	3301	TTCAACAGTC TCGCCGCTAC CGTCTATGCC GACAGTACCG CCGCCCATGC
	3351	CGATATGCAG GGACGCGGCC TGAAAGCCGT ATCGGACGGG TTGGACCACA
25	3401	ACGGCACCGG TCTGCGCGTC ATCGCGCAAACCCAAACAGGA CGGTGGAACG
	3451	TGGGAACAGG GCGGTGTTGA AGGCAAAATG CGCGGCAGTA CCCAAACCGT
	3501	CGGCATTGCC GCGAAAACCG GCGAAAATAC GACAGCAGCC GCCACACTGG
	3551	GCATGGGACG CAGCACATGG AGCGAAAACA GTGCAAATGC AAAAACGAC
	3601	AGCATTAGTC TGTTCGAGG CATAACGGCAC GATGCGGGCG ATATCGGCTA
	3651	TCTCAAAGGC CTGTTCTCCT ACGGACGCTA CAAAACAGC ATCAGCCGCA
30	3701	GCACCGGTGC GGACGAACAT GCGGAAGGCA GCGTCAACCG CACGCTGATG
	3751	CAGCTGGCG CACTGGGGCG TGTCACGTT CGCTTGCG CAACGGGAGA
	3801	TTTGACGGTC GAAGGCAGGTC TGCGCTACGA CCTGCTAAA CAGGATGCAT
	3851	TCGCCGAAAA AGGCAGTGT TTGGGCTGGA GCGGCAACAG CCTCACTGAA
	3901	GGCACGCTGG TCGGACTCGC GGGTCTGAAG CTGTCGCAAC CCTTGAGCGA
	3951	TAAAGCCGTC CTGTTGCAA CGGGGGCGT GGAACCGCAG CTGAACGGAC
35	4001	GCGACTACAC GGTAACGGGC GGCTTACCG GCGCGACTGC AGCAACCGGC
	4051	AAGACGGGG CACCGCAATAT GCCGCACACC CGTCTGGTT CGGGCCTGGG
	4101	CGCGGATGTC GAATTGGCA ACGGCTGGAA CGGCTGGCA CGTTACAGCT
	4151	ACGCCGGTTC CAAACAGTAC GGCAACCACA GCGGACGAGT CGCGTAGGC
	4201	TACCGTTCT GACTCGAG

40

45

50

55

1 MKHFPSKVLT TAILATFCSG ALAATNDDDV KKAATV ріАА АYNNNGQEING  
 51 FKAGETIYDI DEDGTITKKD ATAADVEADD FKGLGLKKVV TNLTKTVNEN  
 101 KQNVDAKVKA AESEIEKLTT KLADTDAALA DTDAALDATT NALNKLGENI  
 151 TTFAEETKTN IVKIDEKLEA VADTVDKHAE AFNDIADSLSD ETNTKADEAV  
 201 KTANEAKQTA EETKQNVDAK VKAAETAAGK AEEAAACTANT AADKAEAVAA  
 251 KVTDIKADIA TNKDNIACKA NSADVYTREE SDSKFRRIDG LNATTEKLDT  
 301 RLASAEKSIA DHDTRLNGLD KTVSDLRKET RQGLAEQAAL SGLFQPYNVG  
 351 GSGGGGTTSAP DFNAGGTGIG SNSRATTAKS AAVSYAGIKN EMCKDRSMLC  
 401 AGRDDVAVTD RDAKINAPPN NLHTGDFPNP NDAYKNLINL KPAIEAGYTG  
 451 RGVEVGIVDT GESVGSISFP ELYGRKEHGY NENYKNYTAY MRKEAPEPDGG  
 501 GKDIЕASFDD EAVIETEAKP TDIRHVKEIG HIDLVSHIIG GRSVDGRPAG  
 551 GIAPDATLHI MNTNDETNE MMVAIRNAW VKLGERGVRI VNNSFGTTSR  
 601 AGTADLFQIA NSEEQYRQAL LDYSGGDKTD EGIRLMQQSD YGNLSYHIRN  
 651 KNMLFIFSTG NDAQAQPNTY ALLPFYEKDA QKGIIITVAGV DRSGEKFKRE  
 701 MYGEPGTEPL EYGSNHCGIT AMWCLSAPYE ASVRFRTRNP IQIAGTSFSA  
 751 PIVTGTAALL LOKYPWMSND NLRTTLLTTA QDIGAVGVDS KFGWGLLDAG  
 801 KAMNGPASFP FGDFTADTKG TSDIAYSFRN DISGTTGLIK KGGSQLQLHG  
 851 NNTYTGKTI EGGSVLVLYGN NKSDMRVETK GALIYNGAAS GGSLNSDGV  
 901 YLADTDQSGA NETVHIKGSN QLDGKGTLYT RLGKLLKVDG TAIIGGKLYM  
 951 SARGKGAGYL NSTGRRVPFL SAAKIGQDYS FFTNIETDGG LLASLDSVEK  
 1001 TAGSEGDTLS YYVRRGNAAR TASAAAHSAP AGLKHAVEQG GSNLENLMVE  
 1051 LDASESSATP ETVETAAADR TDMPGIRPYG ATFRAAAAVQ HANAADGVRI  
 1101 FNLSLAATVYA DSTAAHADMQ GRRLKAVSDG LDHNGTGLRV IAQTQQDGTT  
  
 25 1151 WEQGGVEGKM RGSTQTVGIA AKTGENTTAA ATLCMGRSTW SENSANAKTD  
 1201 SISLFAGIRH DAGDIGYLKG LFSYGRYKNS ISRSTGADEH AEGSVNGTLM  
 1251 QLGALGGVNV PFAATGDLTV EGGLRYDLIK QDAFAEKGSA LGWSGNSLTE  
 1301 GTLVGLAGLK LSQPLSDKAV LFATAGVERD LNGRDYTVTG GFTGATAATG  
 1351 KTGARNMPHT RLVAGLGADV EFGNGWNGLA RYSYAGSKQY GNHSGRVGVG  
 30 1401 YRF\*

[0279] It will be understood that the invention has been described by way of example only and modifications may be made whilst remaining within the scope and spirit of the invention. For instance, the use of proteins from other strains is envisaged [e.g. see WO00/66741 for polymorphic sequences for ORF4, ORF40, ORF46, 225, 235, 287, 519, 726, 919 and 953].

## EXPERIMENTAL DETAILS

### FPLC protein purification

[0280] The following table summarises the FPLC protein purification that was used:

Protein	PI	Column	Buffer	pH	Protocol
121.1 untagged	6.23	Mono Q	Tris	8.0	A
128.1 untagged	5.04	Mono Q	Bis-Tris propane	6.5	A
406.1L	7.75	Mono Q	Diethanolamine	9.0	B
576.1L	5.63	Mono Q	Tris	7.5	B
593 untagged	8.79	Mono S	Hepes	7.4	A
726 untagged	4.95	Hi-trap S	Bis-Tris	6.0	A
919 untagged	10.5(-leader)	Mono S	Bicine	8.5	C
919Lorf4	10.4(-leader)	Mono S	Tris	8.0	B
920L	6.92(-leader)	Mono Q	Diethanolamine	8.5	A
953L	7.56(-leader)	Mono S	MES	6.6	D

(continued)

Protein	PI	Column	Buffer	pH	Protocol
982untagged	4.73	Mono Q	Bis-Tris propane	6.5	A
919-287	6.58	Hi-trap Q	Tris	8.0	A
953-287	4.92	Mono Q	Bis-Tris propane	6.2	A

[0281] Buffer solutions included 20-120 mM NaCl, 5.0 mg/ml CHAPS and 10% v/v glycerol. The dialysate was centrifuged at 13000g for 20 min and applied to either a mono Q or mono S FPLC ion-exchange resin. Buffer and ion exchange resins were chosen according to the pl of the protein of interest and the recommendations of the FPLC protocol manual [Pharmacia: *FPLC Ion Exchange and Chromatofocussing; Principles and Methods*. Pharmacia Publication]. Proteins were eluted using a step-wise NaCl gradient. Purification was analysed by SDS-PAGE and protein concentration determined by the Bradford method.

[0282] The letter in the 'protocol' column refers to the following:

[0283] **FPLC-A:** Clones 121.1, 128.1, 593, 726, 982, periplasmic protein 920L and hybrid proteins 919-287, 953-287 were purified from the soluble fraction of *E.coli* obtained after disruption of the cells. Single colonies harbouring the plasmid of interest were grown overnight at 37°C in 20 ml of LB/Amp (100 µg/ml) liquid culture. Bacteria were diluted 1:30 in 1.0 L of fresh medium and grown at either 30°C or 37°C until the OD<sub>550</sub> reached 0.6-0.8. Expression of recombinant protein was induced with IPTG at a final concentration of 1.0 mM. After incubation for 3 hours, bacteria were harvested by centrifugation at 8000g for 15 minutes at 4°C. When necessary cells were stored at -20°C. All subsequent procedures were performed on ice or at 4°C. For cytosolic proteins (121.1, 128.1, 593, 726 and 982) and periplasmic protein 920L, bacteria were resuspended in 25 ml of PBS containing complete protease inhibitor (Boehringer-Mannheim). Cells were lysed by sonication using a Branson Sonifier 450. Disrupted cells were centrifuged at 8000g for 30 min to sediment unbroken cells and inclusion bodies and the supernatant taken to 35% v/v saturation by the addition of 3.9 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>. The precipitate was sedimented at 8000g for 30 minutes. The supernatant was taken to 70% v/v saturation by the addition of 3.9 M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> and the precipitate collected as above. Pellets containing the protein of interest were identified by SDS-PAGE and dialysed against the appropriate ion-exchange buffer (see below) for 6 hours or overnight. The periplasmic fraction from *E.coli* expressing 953L was prepared according to the protocol of Evans et. al. [Infect.Immun. (1974), 10:1010-1017] and dialysed against the appropriate ion-exchange buffer. Buffer and ion exchange resin were chosen according to the pl of the protein of interest and the recommendations of the FPLC protocol manual (Pharmacia). Buffer solutions included 20 mM NaCl, and 10% (v/v) glycerol. The dialysate was centrifuged at 13000g for 20 min and applied to either a mono Q or mono S FPLC ion-exchange resin. Buffer and ion exchange resin were chosen according to the pl of the protein of interest and the recommendations of the FPLC protocol manual (Pharmacia). Proteins were eluted from the ion-exchange resin using either step-wise or continuous NaCl gradients. Purification was analysed by SDS-PAGE and protein concentration determined by Bradford method. Cleavage of the leader peptide of periplasmic proteins was demonstrated by sequencing the NH<sub>2</sub>-terminus (see below).

[0284] **FPLC-B:** These proteins were purified from the membrane fraction of *E.coli*. Single colonies harbouring the plasmid of interest were grown overnight at 37°C in 20 ml of LB/Amp (100 µg/ml) liquid culture. Bacteria were diluted 1:30 in 1.0 L of fresh medium. Clones 406.1L and 919LOrf4 were grown at 30°C and Orf25L and 576.1L at 37°C until the OD<sub>550</sub> reached 0.6-0.8. In the case of 919LOrf4, growth at 30°C was essential since expression of recombinant protein at 37°C resulted in lysis of the cells. Expression of recombinant protein was induced with IPTG at a final concentration of 1.0 mM. After incubation for 3 hours, bacteria were harvested by centrifugation at 8000g for 15 minutes at 4°C. When necessary cells were stored at -20 °C. All subsequent procedures were performed at 4°C. Bacteria were resuspended in 25 ml of PBS containing complete protease inhibitor (Boehringer-Mannheim) and lysed by osmotic shock with 2-3 passages through a French Press. Unbroken cells were removed by centrifugation at 5000g for 15 min and membranes precipitated by centrifugation at 100000g (Beckman Ti50, 38000rpm) for 45 minutes. A Dounce homogenizer was used to re-suspend the membrane pellet in 7.5 ml of 20 mM Tris-HCl (pH 8.0), 1.0 M NaCl and complete protease inhibitor. The suspension was mixed for 2-4 hours, centrifuged at 100000g for 45 min and the pellet resuspended in 7.5 ml of 20mM Tris-HCl (pH 8.0), 1.0M NaCl, 5.0mg/ml CHAPS, 10% (v/v) glycerol and complete protease inhibitor. The solution was mixed overnight, centrifuged at 100000g for 45 minutes and the supernatant dialysed for 6 hours against an appropriately selected buffer. In the case of Orf25.L, the pellet obtained after CHAPS extraction was found to contain the recombinant protein. This fraction, without further purification, was used to immunise mice.

[0285] **FPLC-C:** Identical to FPLC-A, but purification was from the soluble fraction obtained after permeabilising *E.coli* with polymyxin B, rather than after cell disruption.

[0286] **FPLC-D:** A single colony harbouring the plasmid of interest was grown overnight at 37°C in 20 ml of LB/Amp (100 µg/ml) liquid culture. Bacteria were diluted 1:30 in 1.0 L of fresh medium and grown at 30°C until the OD<sub>550</sub> reached

0.6-0.8. Expression of recombinant protein was induced with IPTG at a final concentration of 1.0mM. After incubation for 3 hours, bacteria were harvested by centrifugation at 8000g for 15 minutes at 4°C. When necessary cells were stored at -20 °C. All subsequent procedures were performed on ice or at 4°C. Cells were resuspended in 20mM Bicine (pH 8.5), 20mM NaCl, 10% (v/v) glycerol, complete protease inhibitor (Boehringer-Mannheim) and disrupted using a Branson Sonifier 450. The sonicate was centrifuged at 8000g for 30 min to sediment unbroken cells and inclusion bodies. The recombinant protein was precipitated from solution between 35% v/v and 70% v/v saturation by the addition of 3.9M  $(\text{NH}_4)_2\text{SO}_4$ . The precipitate was sedimented at 8000g for 30 minutes, resuspended in 20 mM Bicine (pH 8.5), 20 mM NaCl, 10% (v/v) glycerol and dialysed against this buffer for 6 hours or overnight. The dialysate was centrifuged at 13000g for 20 min and applied to the FPLC resin. The protein was eluted from the column using a step-wise NaCl gradients. Purification was analysed by SDS-PAGE and protein concentration determined by Bradford method.

#### **Cloning strategy and oligonucleotide design**

[0287] Genes coding for antigens of interest were amplified by PCR, using oligonucleotides designed on the basis of the genomic sequence of *N. meningitidis* B MC58. Genomic DNA from strain 2996 was always used as a template in PCR reactions, unless otherwise specified, and the amplified fragments were cloned in the expression vector pET21 b+ (Novagen) to express the protein as C-terminal His-tagged product, or in pET-24b+(Novagen) to express the protein in 'untagged' form (e.g. ΔG 287K).

[0288] Where a protein was expressed without a fusion partner and with its own leader peptide (if present), amplification of the open reading frame (ATG to STOP codons) was performed.

[0289] Where a protein was expressed in 'untagged' form, the leader peptide was omitted by designing the 5'-end amplification primer downstream from the predicted leader sequence.

[0290] The melting temperature of the primers used in PCR depended on the number and type of hybridising nucleotides in the whole primer, and was determined using the formulae:

$$T_{m1} = 4 (\text{G+C}) + 2 (\text{A+T}) \quad (\text{tail excluded})$$

$$T_{m2} = 64.9 + 0.41 (\%) \text{GC} - 600/N \quad (\text{whole primer})$$

[0291] The melting temperatures of the selected oligonucleotides were usually 65-70°C for the whole oligo and 50-60°C for the hybridising region alone.

[0292] Oligonucleotides were synthesised using a Perkin Elmer 394 DNA/RNA Synthesizer, eluted from the columns in 2.0ml NH<sub>4</sub>OH, and deprotected by 5 hours incubation at 56°C. The oligos were precipitated by addition of 0.3M Na-Acetate and 2 volumes ethanol. The samples were centrifuged and the pellets resuspended in water.

40

45

50

55

		Sequences	Restriction site
5	<b>Orf1L</b>	Fwd CGCGGATCCGCTAGC-AAAACAACCGACAAACGG	NheI
		Rev CCCG <u>CTCGAG</u> -TTACCAGCGGTAGCCTA	Xhol
10	<b>Orf1L</b>	Fwd CTAGCTAGC-GGACACACTTATTTCGGCATC	NheI
		Rev CCCG <u>CTCGAG</u> - TTACCAGCGGTAGCCTAATTG	Xhol
15	<b>Orf1LOmpA</b>	Fwd	NdeI-(NheI)
		Rev CCCG <u>CTCGAG</u> -	Xhol
20	<b>Orf4L</b>	Fwd CGCGGATCCCATATG-AAAACCTTCTTCAAAACC	NdeI
		Rev CCCG <u>CTCGAG</u> -TTATTGGCTGCGCCTTC	Xhol
25	<b>Orf7-1L</b>	Fwd GCGGC <u>ATTAAT</u> -ATGTTGAGAAAATTGTTGAAATGG	Asel
		Rev GCGGC <u>CTCGAG</u> -TTATTTTTCAAAATATATTGC	Xhol
30	<b>Orf9-1L</b>	Fwd GCGGC <u>CATATG</u> -TTACCTAACCGTTCAAAATGT	NdeI
		Rev GCGGC <u>CTCGAG</u> -TTATTCCGAGGTTTCGGG	Xhol
35	<b>Orf23L</b>	Fwd CGCGGATCCCATATG-ACACGCTTCAAATATT	NdeI
		Rev CCCG <u>CTCGAG</u> -TTATTAAACCGATAGGTAAA	Xhol
40	<b>Orf25-1 His</b>	Fwd CGCGGATCCCATATG-GGCAGGGAAGAACCGC	NdeI
		Rev GCCCAAG <u>CTT</u> -ATCGATGGAATAGCCGCC	HindIII
45	<b>Orf29-1 b-His (MC58)</b>	Fwd CGCGGATCCGCTAGC-AACGGTTGGATGCCCG	NheI
		Rev CCCG <u>CTCGAG</u> -TTTGCTTAAGTCCCTGATAT CCCG <u>CTCGAG</u> -ATTCCCACCTGCCATC	Xhol
50	<b>Orf29-1 b-L (MC58)</b>	Fwd CGCGGATCCGCTAGC-ATGAATTGCTATTCAAAAT	NheI
		Rev CCCG <u>CTCGAG</u> -TTAATTCCCACCTGCCATC	Xhol
55	<b>Orf29-1 c-His (MC58)</b>	Fwd CGCGGATCCGCTAGC-ATGAATTGCTATTCAAAAT	NheI
		Rev CCCG <u>CTCGAG</u> -TTGGACGATGCCGCCGA	Xhol
60	<b>Orf29-1 c-L (MC58)</b>	Fwd CGCGGATCCGCTAGC-ATGAATTGCTATTCAAAAT	NheI
		Rev CCCG <u>CTCGAG</u> -TTATTGGACGATGCCGCC	Xhol
65	<b>Orf25L</b>	Fwd CGCGGATCCCATATG-TATCGCAAACGTATTGC	NdeI
		Rev CCCG <u>CTCGAG</u> -CTAATCGATGGAATAGCC	Xhol
70	<b>Orf37L</b>	Fwd CGCGGATCCCATATG-AAACAGACAGTCAAATG	NdeI
		Rev CCCG <u>CTCGAG</u> -TCAATAACCGCCTTCAG	Xhol
75	<b>Orf38L</b>	Fwd CGCGGATCCCATATG- TTACGTTGACTGCTTAGCCGTATGCACC	NdeI
		Rev CCCG <u>CTCGAG</u> - TTATTTGCCGCGTAAAAGCGCTGGCAAC	Xhol
80	<b>Orf40L</b>	Fwd CGCGGATCCCATATG-AACAAATATACCGCAT	NdeI
		Rev CCCG <u>CTCGAG</u> -TTACCACTGATAACCGAC	Xhol
85	<b>Orf40.2-His</b>	Fwd CGCGGATCCCATATG-ACCGATGACGACGGATTAT	NdeI
		Rev GCCCAAG <u>CTT</u> -CCACTGATAACCGACAGA	HindIII
90	<b>Orf40.2L</b>	Fwd CGCGGATCCCATATG-AACAAATATACCGCAT	NdeI
		Rev GCCCAAG <u>CTT</u> -TTACCACTGATAACCGAC	HindIII
95	<b>Orf46-2L</b>	Fwd GGGATTCCATATG-GGCATTCCCGCAAAATATC	NdeI
		Rev CCCG <u>CTCGAG</u> -TTATTTACTCCTATAACGAGGTCTCTAAC	Xhol
100	<b>Orf46-2</b>	Fwd GGGATTCCATATG-TCAGATTGGCAAACGATTCTT	NdeI
		Rev CCCG <u>CTCGAG</u> -TTATTTACTCCTATAACGAGGTCTCTAAC	Xhol
105	<b>Orf46.1L</b>	Fwd GGGATTCCATATG-GGCATTCCCGCAAAATATC	NdeI

		Rev	<u>CCCGCTCGAG-TTACGTATCATATTCACGTGC</u>	Xhol
5	<b>orf46. (His-GST)</b>	Fwd	<u>GGGAATTCCATATGACGTGAAATATGATACGAAG</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAGTTACTCCTATAACGAGGTCTCTAAC</u>	Xhol
10	<b>orf46.1-His</b>	Fwd	<u>GGGAATTCCATATGTCAGATTGGCAAACGATTCTT</u>	NdeI
		Rev	<u>CCCGCTCGAGCGTATCATATTCACGTGC</u>	Xhol
15	<b>orf46.2-His</b>	Fwd	<u>GGGAATTCCATATGTCAGATTGGCAAACGATTCTT</u>	NdeI
		Rev	<u>CCCGCTCGAGTTACTCCTATAACGAGGTCTCTAAC</u>	Xhol
20	<b>Orf65-1-(His/GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-CAAAATGCGTTAAAATCCC</u>	BamHI-NdeI
		Rev	<u>CGCGGATCCCATATG-AACAAATATACCGCAT CCCGCTCGAG -TTGCTTCGATAGAACGG</u>	Xhol
	<b>Orf72-1L</b>	Fwd	<u>GCGGCCATATG-GTCATAAAATACAAATTGAA</u>	NdeI
		Rev	<u>GCGGCCTCGAG-TTAGCCTGAGACCTTGCAAATT</u>	Xhol
25	<b>Orf76-1L</b>	Fwd	<u>GCGGCCATATG-AAACAGAAAAAACCGCTG</u>	NdeI
		Rev	<u>GCGGCCTCGAG-TTACGGTTGACACCGTTTC</u>	Xhol
	<b>Orf83.1L</b>	Fwd	<u>CGCGGATCCCATATG-AAAACCTGCTCCTC</u>	NdeI
		Rev	<u>CCCGCTCGAG-TTATCCTCCTTGCAGC</u>	Xhol
30	<b>Orf85-2L</b>	Fwd	<u>GCGGCCATATG-GCAAAATGATGAAATGGG</u>	NdeI
		Rev	<u>GCGGCCTCGAG-TTATCGGCGGGCGGGCC</u>	Xhol
	<b>Orf91L (MC58)</b>	Fwd	<u>GCGGCCATATGAAAAAATCCTCCCTCATCA</u>	NdeI
		Rev	<u>GCGGCCTCGAGTTATTGCCGCCGTTTTGGC</u>	Xhol
35	<b>Orf91-His(MC58)</b>	Fwd	<u>GCGGCCATATGCCCTGCGACGCGGTAAG</u>	NdeI
		Rev	<u>GCGGCCTCGAGTTGCCGCCGTTTTGGCTTC</u>	Xhol
	<b>Orf97-1L</b>	Fwd	<u>GCGGCCATATG-AAACACATACTCCCCCTGA</u>	NdeI
		Rev	<u>GCGGCCTCGAG-TTATCGCTACGGTTTTG</u>	Xhol
40	<b>Orf119L (MC58)</b>	Fwd	<u>GCGGCCATATGATTACATCGTACTGTTTC</u>	NdeI
		Rev	<u>GCGGCCTCGAGTTAGGAGAACAGCGCAATGC</u>	Xhol
	<b>Orf119-His(MC58)</b>	Fwd	<u>GCGGCCATATGTACAACATGTATCAGGAAAC</u>	NdeI
		Rev	<u>GCGGCCTCGAGGGAGAACAGCGCAATCGG</u>	Xhol
45	<b>Orf137.1 (His-GST) (MC58)</b>	Fwd	<u>CGCGGATCCGCTAGCTCGGGCACGGCGGG</u>	BamHI-NheI
		Rec	<u>CCCGCTCGAGATAACGGTATGCCGCCAG</u>	Xhol
	<b>Orf143-1L</b>	Fwd	<u>CGCGGATCCCATATG-GAATCAACACTTCAC</u>	NdeI
		Rev	<u>CCCGCTCGAG-TTACACCGCGTTGCTGT</u>	Xhol
50	<b>008</b>	Fwd	<u>CGCGGATCCCATATG-AACAAACAGACATTG</u>	NdeI
		Rev	<u>CCCGCTCGAG-TTACCTGTCCGGTAAAG</u>	Xhol
	<b>050-1(48)</b>	Fwd	<u>CGCGGATCCGCTAGC-ACCGTCATCAAACAGGAA</u>	NheI
		Rev	<u>CCCGCTCGAG-TCAAGATTGACGGGGA</u>	Xhol
55	<b>105</b>	Fwd	<u>CGCGGATCCCATATG-TCCGCAAACGAATACG</u>	NdeI
		Rev	<u>CCCGCTCGAG-TCAGTGTCTGCCAGTT</u>	Xhol
	<b>111L</b>	Fwd	<u>CGCGGATCCCATATG-CCGTCTGAAACACG</u>	NdeI
		Rev	<u>CCCGCTCGAG-TTAGCGGAGCAGTTTTTC</u>	Xhol
55	<b>117-1</b>	Fwd	<u>CGCGGATCCCATATG-ACGCCATCAGCC</u>	NdeI
		Rev	<u>CCCGCTCGAG-TTAAAGCCGGTAACGC</u>	Xhol
	<b>121-1</b>	Fwd	<u>CGGGCCATATG-GAAACACAGCTTACATCGG</u>	NdeI
		Rev	<u>GCGGCCTCGAG-TCAATAATAATACCCCGCG</u>	Xhol

	<b>122-1</b>	Fwd	<u>GCGGCCATATG</u> -ATTAATCCGCAATATCC	NdeI
		Rev	<u>GCGGCCTCGAG</u> -TTAAATCTTGGTAGATTGGATTGG	Xhol
5	<b>128-1</b>	Fwd	<u>GCGGCCATATG</u> -ACTGACAACGCACTGCTCC	NdeI
		Rev	<u>GCGGCCTCGAG</u> -TCAGACCGCGTGTGCAAAC	Xhol
10	<b>148</b>	Fwd	<u>CGCGGATCCCATATG</u> -GCGTTAAAAACATCAA	NdeI
		Rev	<u>CCCGCTCGAG</u> -TCAGCCCTCATACAGC	Xhol
15	<b>149.1L (MC58)</b>	Fwd	<u>GCGGCATTAATGGCACAAACTACACTCAAACC</u>	Ascl
		Rev	<u>GCGGCCTCGAGTTAAAACCTTCACGTTACGCCG</u>	Xhol
20	<b>149.1-His(MC58)</b>	Fwd	<u>GCGGCATTAATGCATGAAACTGAGCAATCGGTGG</u>	Ascl
		Rev	<u>GCGGCCTCGAGAAACTTCACGTTACGCCGCCG</u>	Xhol
25	<b>205 (His-GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATGGCAAATCCGAAAATACG</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAGATAATGGCGGCCGCG</u>	Xhol
30	<b>206L</b>	Fwd	<u>CGCGGATCCCATATG</u> -TTTCCCCCGACAA	NdeI
		Rev	<u>CCCGCTCGAG</u> -TCATTCTGTAAAAAAAGTATG	Xhol
35	<b>214 (His-GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATGCTCAAAGCGACAGCAG</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAGTTCGGATTTTGCGTACTC</u>	Xhol
40	<b>216</b>	Fwd	<u>CGCGGATCCCATATG</u> -GCAATGGCAGAAAACG	NdeI
		Rev	<u>CCCGCTCGAG</u> -CTATACAATCCGTGCCG	Xhol
45	<b>225-1L</b>	Fwd	<u>CGCGGATCCCATATG</u> -GATTCTTTTCAAACC	NdeI
		Rev	<u>CCCGCTCGAG</u> -TCAGTTCAGAAAGCGGG	Xhol
50	<b>235L</b>	Fwd	<u>CGCGGATCCCATATG</u> -AACCTTTGATTTAGG	NdeI
		Rev	<u>CCCGCTCGAG</u> -TTATTGGGCTGCTCTTC	Xhol
55	<b>243</b>	Fwd	<u>CGCGGATCCCATATG</u> -GTAATCGTCTGGTTG	NdeI
		Rev	<u>CCCGCTCGAG</u> -CTACGACTTGGTTACCG	Xhol
60	<b>247-1L</b>	Fwd	<u>CGGGCCATATG</u> -AGACGTAAAATGCTAAAGCTAC	NdeI
		Rev	<u>GCGGCCTCGAG</u> -TCAAAGTGTCTGTTGCGC	Xhol
65	<b>264-His</b>	Fwd	<u>GCGGCCATATG</u> -TTGACTTTAACCGAAAAA	NdeI
		Rev	<u>GCGGCCTCGAG</u> -GCCGGCGGTCAATACGCCGAA	Xhol
70	<b>270 (His-GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATGGCGCAATCGGATTGAC</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAGTTGGCGGTAAATGCCG</u>	Xhol
75	<b>274L</b>	Fwd	<u>CGGGCCATATG</u> -GCGGGGCCGATTTTGT	NdeI
		Rev	<u>GCGGCCTCGAG</u> -TTATTGCTTCAGTATTATTG	Xhol
80	<b>283L</b>	Fwd	<u>CGGGCCATATG</u> -AACTTGCTTATCCGTCA	NdeI
		Rev	<u>CGGGCCTCGAG</u> -TTAACGGCAGTATTGTTAC	Xhol
85	<b>285-His</b>	Fwd	<u>CGCGGATCCCATATGGGTTGCGCTTCGGC</u>	BamHI
		Rev	<u>GCCCAAGCTTTCCCTTGCCGTTCCG</u>	HindIII
90	<b>286-His (MC58)</b>	Fwd	<u>CGCGGATCCCATATG</u> -GCCGACCTTCCGAAAAA	NdeI
		Rev	<u>CCCGCTCGAG</u> -GAAGCGCGTCCCAAGC	Xhol
95	<b>286L (MC58)</b>	Fwd	<u>CGCGGATCCCATATG</u> -CACGACACCGTAC	NdeI
		Rev	<u>CCCGCTCGAG</u> -TTAGAACGCAGCGTAATCGCAATGG	Xhol
100	<b>287L</b>	Fwd	<u>CTAGCTAGC</u> -TTAACGCAGCGTAATCGCAATGG	NheI
		Rev	<u>CCCGCTCGAG</u> -TCAATCCTGCTTTTGCC	Xhol

	287	Fwd	<u>CTAGCTAGC-GGGGGCGGCGGTGGCG</u>	NheI
		Rev	<u>CCCGCTCGAG-TCAATCCTGCTCTTTTGCC</u>	XhoI
5	<b>287LOrf4</b>	Fwd	<u>CTAGCTAGC-GCTCATCCTCGCCGCC-TGC</u> <u>GGGGGGCGGCGGT</u>	NheI
		Rev	<u>CCCGCTCGAG-TCAATCCTGCTCTTTTGCC</u>	XhoI
10	<b>287-fu</b>	Fwd	<u>CGGGGATCC-GGGGGCGGCGGTGGCG</u>	BamHI
		Rev	<u>CCCGCTCGAG-TCAATCCTGCTCTTTTGCC</u>	XhoI
15	<b>287-His</b>	Fwd	<u>CTAGCTAGC-GGGGGCGGCGGTGGCG</u>	NheI
		Rev	<u>CCCGCTCGAG-ATCCTGCTCTTTTGCC</u> *	XhoI
20	<b>287-His(2996)</b>	Fwd	<u>CTAGCTAGC-TGC</u> <u>GGGGGGCGGCGGTGGCG</u>	NheI
		Rev	<u>CCCGCTCGAG-ATCCTGCTCTTTTGCC</u>	XhoI
25	<b>Δ1 287-His</b>	Fwd	<u>CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC</u> §	NheI
	<b>Δ2 287-His</b>	Fwd	<u>CGCGGATCCGCTAGC-CAAGATATGGCGGCAGT</u> §	NheI
30	<b>Δ3 287-His</b>	Fwd	<u>CGCGGATCCGCTAGC-GCCGAATCCGCAAATCA</u> §	NheI
	<b>Δ4 287-His</b>	Fwd	<u>CGCGCTAGC-GGAAGGGTTGATTGGCTAATGG</u> §	NheI
35	<b>Δ4 287MC58-His</b>	Fwd	<u>CGCGCTAGC-GGAAGGGTTGATTGGCTAATGG</u> §	NheI
	<b>287a-His</b>	Fwd	<u>CGCCATATG-TTTAAACGCAGCGTAATCGC</u>	NdeI
		Rev	<u>CCCGCTCGAG-AAAATTGCTACCGCCATTGCAGG</u>	XhoI
40	<b>287b-His</b>	Fwd	<u>CGCCATATG-GGAAGGGTTGATTGGCTAATGG</u>	NdeI
	<b>287b-2996-His</b>	Rev	<u>CCCGCTCGAG-CTTGTCTTTATAAATGATGACATATTG</u>	XhoI
45	<b>287b-MC58-His</b>	Rev	<u>CCCGCTCGAG-TTATAAAAGATAATATATTGATTGATTCC</u>	XhoI
	<b>287c-2996-His</b>	Fwd	<u>CGCGCTAGC-ATGCCGCTGATTCCCGTCAATC</u> §	NheI
50	<b>'287'untagged'(2996)</b>	Fwd	<u>CTAGCTAGC-GGGGGCGGCGGTGGCG</u>	NheI
		Rev	<u>CCCGCTCGAG-TCAATCCTGCTCTTTTGCC</u>	XhoI
55	<b>ΔG287-His *</b>	Fwd	<u>CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC</u>	NheI
		Rev	<u>CCCGCTCGAG-ATCCTGCTCTTTTGCC</u>	XhoI
55	<b>ΔG287K(2996)</b>	Fwd	<u>CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC</u>	NheI
		Rev	<u>CCCGCTCGAG-TCAATCCTGCTCTTTTGCC</u>	XhoI
55	<b>ΔG 287-L</b>	Fwd	<u>CGCGGATCCGCTAGC-TTGAAACGAGTGTATTGCAATGGCTTGTATTITGCC</u> <u>CTTCAGCGCTGT TCGCCCGATGTTAAATCGGC</u>	NheI
		Rev	<u>CCCGCTCGAG-TCAATCCTGCTCTTTTGCC</u>	XhoI
55	<b>ΔG 287-Orf4L</b>	Fwd	<u>CGCGGATCCGCTAGC-AAAACCTTCTAAACCCCTTCCGCCGCGCACTCGCG</u> <u>CTCATCCTGCCGCCTGC TCGCCCGATGTTAAATCG</u>	NheI
		Rev	<u>CCCGCTCGAG-TCAATCCTGCTCTTTTGCC</u>	XhoI
55	<b>292L</b>	Fwd	<u>CGCGGATCCCATATG-AAAACCAAGTTAACAAA</u>	NdeI
		Rev	<u>CCCGCTCGAG-TTATTGATTGGCGGATGA</u>	XhoI
55	<b>308-1</b>	Fwd	<u>CGCGGATCCCATATG-TTAAATCGGGTATTTATC</u>	NdeI
		Rev	<u>CCCGCTCGAG-TTAAATCCGCCATTCCCTG</u>	XhoI
55	<b>401L</b>	Fwd	<u>GCGGCCATATG-AAATTACAACAATTGGCTG</u>	NdeI
		Rev	<u>GC GG CCTCGAG-TTACCTTACGTTTCAAAG</u>	XhoI
55	<b>406L</b>	Fwd	<u>CGCGGATCCCATATG-CAAGCACGGCTGCT</u>	NdeI
		Rev	<u>CCCGCTCGAG-TCAAGGTTGTCTTGTCTA</u>	XhoI
55	<b>502-1L</b>	Fwd	<u>CGCGGATCCCATATG-ATGAAACCGCACAAC</u>	NdeI
		Rev	<u>CCCGCTCGAG-TCAAGGTTGTCTTGTCTA</u>	XhoI

5	<b>502-A (His-GST)</b>	Fwd	CGCGGATCCCATATGGTAGACCGCGCTTAAGCA	BamHI-NdeI
		Rev	CCCGCTCGAGAGCTGCATGGCGCG	Xhol
10	<b>503-1L</b>	Fwd	CGCGGATCCCATATG-GCACGGTCGTATAC	NdeI
		Rev	CCCGCTCGAG-CTACCGCGCATTCCTG	Xhol
15	<b>519-1L</b>	Fwd	GCGGCCATATG-GAATTTTCAATTATCTTGT	NdeI
		Rev	GCGGCCTCGAG-TTATTGGCGGTTTGCTGC	Xhol
20	<b>525-1L</b>	Fwd	GCGGCCATATG-AAGTATGTCCGGTATTTC	NdeI
		Rev	GCGGCCTCGAG-TTATCGGTTGTGCAACGG	Xhol
25	<b>529-(His/GST) (MC58)</b>	Fwd	CGCGGATCCGCTAGC-TCCGGCAGCAAAACGA	Bam HI-NheI
		Rev	GCCCAAGCTT-ACGCAGTTCGGAATGGAG	HindIII
30	<b>552L</b>	Fwd	GCCGCCATATGTTGAATATTAAACTGAAAACCTTG	NdeI
		Rev	GCCGCCTCGAGTTATTCTGATGCCCTTCCC	Xhol
35	<b>556L</b>	Fwd	GCCGCCATATGGACAATAAGACCAAACGT	NdeI
		Rev	GCCGCCTCGAGTTAACGGTGCAGACGTTTC	Xhol
40	<b>557L</b>	Fwd	CGCGGATCCCATATG-AACAAACTGTTCTTAC	NdeI
		Rev	CCCGCTCGAG-TCATTCCGCCTTCAGAAA	Xhol
45	<b>564ab-(His/GST) (MC58)</b>	Fwd	CGCGGATCCCATATG-CAAGGTATCGTGCAGACAAATCCGCACCT	BamHI-NdeI
		Rev	CCCGCTCGAG-AGCTAATTGTGCTTGGTTGCAGATAGGAGTT	Xhol
50	<b>564abL (MC58)</b>	Fwd	CGCGGATCCCATATG-AACCGCACCCCTGTACAAAGTTGATTTAACAAACATC	NdeI
		Rev	CCCGCTCGAG-TTAAGCTAATTGTGCTTGGTTGCAGATAGGAGTT	Xhol
55	<b>564b-(His/GST)(MC58)</b>	Fwd	CGCGGATCCCATATG-ACGGGAGAAAATCATGCGGTTTCACITCATG	BamHI-NdeI
		Rev	CCCGCTCGAG-AGCTAATTGTGCTTGGTTGCAGATAGGAGTT	Xhol
60	<b>564c-(His/GST)(MC58)</b>	Fwd	CGCGGATCCCATATG-GTTTCAGACGGCCTATACAACCAACATGGTAAATT	BamHI-NdeI
		Rev	CCCGCTCGAG-GCGGTAACTGCCGCTTGCAGTAAATCCGTAA	Xhol
65	<b>564bc-(His/GST)(MC58)</b>	Fwd	CGCGGATCCCATATG-ACGGGAGAAAATCATGCGGTTTCACITCATG	BamHI-NdeI
		Rev	CCCGCTCGAG-GCGGTAACTGCCGCTTGCAGTAAATCCGTAA	Xhol
70	<b>564d-(His/GST)(MC58)</b>	Fwd	CGCGGATCCCATATG-CAAAGCAAAGTCAAAGCAGACCATGCCTCCGTAA	BamHI-NdeI
		Rev	CCCGCTCGAG-TCTTTCTTCAATTATAACTTTAGTAGGTTCAATTTCG GTCCCC	Xhol
75	<b>564cd-(His/GST)(MC58)</b>	Fwd	CGCGGATCCCATATG-GTTTCAGACGGCCTATACAACCAACATGGTAAATT	BamHI-NdeI
		Rev	CCCGCTCGAG-TCTTTCTTCAATTATAACTTTAGTAGGTTCAATTTCG GTCCCC	Xhol
80	<b>570L</b>	Fwd	GCGGCCATATG-ACCCGTTGACCCGCG	NdeI
		Rev	GCGGCCTCGAG-TCAGCGGGCGTTCAATTCTT	Xhol
85	<b>576-1L</b>	Fwd	CGCGGATCCCATATG-AACACCATTTCAAAATC	NdeI
		Rev	CCCGCTCGAG-TTAATTACTTTTGATGTCG	Xhol

	<b>580L</b>	Fwd	GCGGCCATATG-GATTGCCCAAGGTCGG	NdeI
		Rev	GCGGCCTCGAG-CTACACTTCCCCGAAGTGG	Xhol
5	<b>583L</b>	Fwd	CGCGGATCCCATATG-ATAGTTGACCAAAGCC	NdeI
		Rev	CCCGCTCGAG-TTATTTTCGATTTTCGG	Xhol
10	<b>593</b>	Fwd	GCGGCCATATG-CTTGAACGTAAACGGACT	NdeI
		Rev	GCGGCCTCGAG-TCAGCGGAAGCGGACGATT	Xhol
15	<b>650 (His-GST) (MC58)</b>	Fwd	CGCGGATCCCATATGTCCAAACCTAAAACCATCG	BamHI-NdeI
		Rev	CCCGCTCGAGGCTTCCAATCAGTTGACC	Xhol
20	<b>652</b>	Fwd	GCGGCCATATG-AGCGCAATCGTGATATTTTC	NdeI
		Rev	GCGGCCTCGAG-TTATTTGCCAGTTGGTAGAATG	Xhol
25	<b>664L</b>	Fwd	GCGGCCATATG-GTGATAACATCCGCACTACTTC	NdeI
		Rev	GCGGCCTCGAG-TCAAAATCGAGTTTACACCA	Xhol
30	<b>726</b>	Fwd	GCGGCCATATG-ACCATCTATTCAAAACGG	NdeI
		Rev	GCGGCCTCGAG-TCAGCCGATGTTAGCGTCCATT	Xhol
35	<b>741-His(MC58)</b>	Fwd	CGCGGATCCCATATG-AGCAGCGGAGGGGGTG	NdeI
		Rev	CCCGCTCGAG-TTGCTGGCGGCAAGGC	Xhol
40	<b>ΔG741-His(MC58)</b>	Fwd	CGCGGATCCCATATG-GTCGCCGCCGACATCG	NdeI
		Rev	CCCGCTCGAG-TTGCTGGCGGCAAGGC	Xhol
45	<b>686-2-(His/GST) (MC58)</b>	Fwd	CGCGGATCCCATATG-GGCGGTTCGGAAGGCG	BamHI-NdeI
		Rev	CCCGCTCGAG-TTGAAACACTGATGTCTTCCGA	Xhol
50	<b>719-(His/GST) (MC58)</b>	Fwd	CGCGGATCCGCTAGC-AAACTGTCGTTGGTGTAAAC	BamHI-NheI
		Rev	CCCGCTCGAG-TTGACCCGCTCACCG	Xhol
55	<b>730-His (MC58)</b>	Fwd	GCCGCCATATGGCGGACTTGGCGCAAGACCC	NdeI
		Rev	GCCGCCTCGAGATCTCTAACCTGTTAACATGCCG	Xhol
60	<b>730A-His (MC58)</b>	Fwd	GCCGCCATATGGCGGACTTGGCGCAAGACCC	NdeI
		Rev	GCCGCCTCGAGCTCATGCTGTTGCCCAAGC	Xhol
65	<b>730B-His (MC58)</b>	Fwd	GCCGCCATATGGCGGACTTGGCGCAAGACCC	NdeI
		Rev	GCCGCCTCGAGAAAATCCCCGCTAACCGCAG	Xhol
70	<b>741-His (MC58)</b>	Fwd	CGCGGATCCCATATG-AGCAGCGGAGGGGGTG	NdeI
		Rev	CCCGCTCGAG-TTGCTGGCGGCAAGGC	Xhol
75	<b>ΔG741-His (MC58)</b>	Fwd	CGCGGATCCCATATG-GTCGCCGCCGACATCG	NdeI
		Rev	CCCGCTCGAG-TTGCTGGCGGCAAGGC	Xhol
80	<b>743 (His-GST)</b>	Fwd	CGCGGATCCCATATGGACGGTGTGCTGCTGTT	BamHI-NdeI
		Rev	CCCGCTCGAGCTTACGGATCAAATTGACG	Xhol
85	<b>757 (His-GST) (MC58)</b>	Fwd	CGCGGATCCCATATGGCAGCCAATCTGAAGAA	BamHI-NdeI
		Rev	CCCGCTCGAGCTCAGCTTGTGCCGTCAA	Xhol
90	<b>759-His/GST (MC58)</b>	Fwd	CGCGGATCCGCTAGC-TACTCATCCATTGTCCGC	BamHI-NheI
		Rev	CCCGCTCGAG-CCAGTTGTAGCCTATTTG	Xhol
95	<b>759L (MC58)</b>	Fwd	CGCGGATCCGCTAGC-ATGCGCTTCACACACAC	NheI
		Rev	CCCGCTCGAG-TTACCAAGTTGTAGCCTATT	Xhol
100	<b>760-His</b>	Fwd	GCCGCCATATGGCACAAACGGAAGGTTGGAA	NdeI
		Rev	GCCGCCTCGAGAAAATGTAACGCAAGGTTGCCGTC	Xhol
105	<b>769-His (MC58)</b>	Fwd	GCGGCCATATGGAAGAACACCGCGCAACCG	NdeI

		Rev	GCGGCCTCGAGGAACGTTTATTAAACTCGAC	Xhol
5	907L	Fwd	GCGGCCATATG-AGAAAACCGACCGATAACCCTA	NdeI
		Rev	GCGGC <u>CTCGAG</u> -TCAACGCCACTGCCAGCGGTTG	Xhol
10	911L	Fwd	CGCGGATCCC <u>ATATG</u> -AAGAAGAACATATTGAATTTGGGTCGGACTG	NdeI
		Rev	CCCG <u>CTCGAG</u> -TTATTCGGCGGTTTTCCGCATTGCCG	Xhol
15	911LOmpA	Fwd	GGGAATTCC <u>CATATG</u> AAAAAGACAGCTATCGCGATTGCA GTGGCACTGGCTGGTTTCGCTACCGTAGCGCAGGCC <u>GC</u> <u>TAGC</u> -GCTTCCCGTGGCCGGCGGTGC	NdeI-(NheI)
		Rev	CCCG <u>CTCGAG</u> -TTATTCGGCGGTTTTCCGCATTGCCG	Xhol
20	911LPelB	Fwd	CATGCC <u>ATGG</u> -CTTCCCGCGTGGCCGGCGGTGC	Ncol
		Rev	CCCG <u>CTCGAG</u> -TTATTCGGCGGTTTTCCGCATTGCCG	Xhol
25	913-His/GST (MC58)	Fwd	CGCGGATCCC <u>CATATG</u> -TTGCCGAAACCCGCC	BamHI-NdeI
		Rev	CCCG <u>CTCGAG</u> -AGGTTGTGTTCCAGGTTG	Xhol
30	913L (MC58)	Fwd	CGCGGATCCC <u>CATATG</u> -AAAAAAACCGCCTATG	NdeI
		Rev	CCCG <u>CTCGAG</u> -TTAAGGTTGTGTTCCAGG	Xhol
35	919L	Fwd	CGCGGATCCC <u>CATATG</u> -AAAAAAATACCTATTCCGC	NdeI
		Rev	CCCG <u>CTCGAG</u> -TTACGGGCGGTATTCCGG	Xhol
40	919	Fwd	CGCGGATCCC <u>CATATG</u> -CAAAGCAAGAGCATCCAAA	NdeI
		Rev	CCCG <u>CTCGAG</u> -TTACGGGCGGTATTCCGG	Xhol
45	919L Orf4	Fwd	GGGAATTCC <u>CATATG</u> AAAACCTCTTCAAAACCCTTCCG CCGCC <u>CGCTAGCGCTCATCCTCGCCGCC</u> TGCAAAGCAAGAGCATC	NdeI-(NheI)
		Rev	CCCG <u>CTCGAG</u> -TTACGGGCGGTATTCCGGTACACCG	Xhol
50	(919)-287fusion	Fwd	CGCGGATCC <u>GTGAC</u> -TGTGGGGCGGGCGGTGGC	Sall
		Rev	CCCG <u>CTCGAG</u> -TCAATCCTGCTTTTGCC	Xhol
55	920-1L	Fwd	GCGGCC <u>CATATG</u> -AAGAAAACATTGACACTGC	NdeI
		Rev	GCGGC <u>CTCGAG</u> -TTAATGGTGCAGATGACCGAT	Xhol
60	925-His/GST (MC58) <sup>GATE</sup>	Fwd	ggggacaagtttacaaaaaagcaggctTGCGGAAGGATGCCGG	<i>attB1</i>
		Rev	ggggaccactttacaaagaagctgggtCTAAAGCAACAATGCCGG	<i>attB2</i>
65	926L	Fwd	CGCGGATCCC <u>CATATG</u> -AAACACACCGTATCC	NdeI
		Rev	CCCG <u>CTCGAG</u> -TTATCTCGTGCAGGCC	Xhol
70	927-2-(His/GST) (MC58)	Fwd	CGCGGATCCC <u>CATATG</u> -AGCCCCGCCGATT	BamHI-NdeI
		Rev	CCCG <u>CTCGAG</u> -TTTTGTGCGGTCAAGGCG	Xhol
75	932-His/GST (MC58) <sup>GATE</sup>	Fwd	ggggacaagtttacaaaaaagcaggctTGTCGTTGGGGATTAA ACCAAAACCAATC	<i>attB1</i>
		Rev	CGCGGATCCC <u>CATATG</u> -GGCGGATGCAGGCC	BamHI-NdeI
80	935 (His-GST) (MC58)	For	CCCG <u>CTCGAG</u> AAACCGCCAATCCGCC	Xhol
		Rev	ggggaccactttacaaagaagctgggtTCATTTGTTTCCTTCT CGAGGCCATT	<i>attB2</i>
85	936-1L	Fwd	CGCGGATCCC <u>CATATG</u> -AAACCAAAACCGCAC	NdeI
		Rev	CCCG <u>CTCGAG</u> -TCAGCGTTGGACGTAGT	Xhol
90	953L	Fwd	GGGAATTCC <u>CATATG</u> -AAAAAAATCATCTCGCCG	NdeI
		Rev	CCCG <u>CTCGAG</u> -TTATTGTTGGCTGCCTCGAT	Xhol
95	953-fu	Fwd	GGGAATTCC <u>CATATG</u> -GCCACCTACAAAGTGGACG	NdeI
		Rev	CGGGGATCC-TTGTGGCTGCCTCGATTG	BamHI

5	954 (His-GST) (MC58)	Fwd	CGCGGATCCCATATGCAAGAACAAATCGCAGAAAG	BamHI-NdeI	
		Rev	CCCGCTCGAGTTTTCGGCAAATTGGCTT	XhoI	
10	958-His/GST (MC58) GATE	Fwd	ggggacaagtttacaaaaaaaggcaggctGCCGATGCCGTTGCGG	<i>attB1</i>	
	961L	Rev	ggggaccacttgacaaagactgggTCAGGGTCGTTGTTGCGG	<i>attB2</i>	
15		Fwd	CGCGGATCCCATATG-AAACACTTCCATCC	NdeI	
		Rev	CCCGCTCGAG-TTACCACTCGTAATTGAC	XhoI	
20	961	Fwd	CGCGGATCCCATATG-GCCACAAGCGACGAC	NdeI	
	Rev	CCCGCTCGAG-TTACCACTCGTAATTGAC	XhoI		
25	961 c (His/GST)	Fwd	CGCGGATCCCATATG-GCCACAAACGACG	BamHI-NdeI	
		Rev	CCCGCTCGAG-ACCCACGTTGTAAGGTTG	XhoI	
30	961 c-(His/GST) (MC58)	Fwd	CGCGGATCCCATATG-GCCACAAGCGACGACGA	BamHI-NdeI	
		Rev	CCCGCTCGAG-ACCCACGTTGTAAGGTTG	XhoI	
35	961 c-L	Fwd	CGCGGATCCCATATG-ATGAAACACTTCCATCC	NdeI	
		Rev	CCCGCTCGAG-TTAACCCACGTTGTAAGGT	XhoI	
40	961 c-L (MC58)	Fwd	CGCGGATCCCATATG-ATGAAACACTTCCATCC	NdeI	
		Rev	CCCGCTCGAG-TTAACCCACGTTGTAAGGT	XhoI	
45	961 d (His/GST)	Fwd	CGCGGATCCCATATG-GCCACAAACGACG	BamHI-NdeI	
		Rev	CCCGCTCGAG-GTCTGACACTGTTTATCC	XhoI	
50	961 Δ1-L	Fwd	CGCGGATCCCATATG-ATGAAACACTTCCATCC	NdeI	
		Rev	CCCGCTCGAG-TTATGCTTGGCGGCAAAG	XhoI	
55	fu 961-...	Fwd	CGCGGATCCCATATG- GCCACAAACGACGAC	NdeI	
		Rev	CGCGGATCC-CCACTCGTAATTGACGCC	BamHI	
60	fu 961-... (MC58)	Fwd	CGCGGATCCCATATG-GCCACAAGCGACGAC	NdeI	
		Rev	CGCGGATCC-CCACTCGTAATTGACGCC	BamHI	
65	fu 961 c -...	Fwd	CGCGGATCCCATATG-GCCACAAACGACGAC	NdeI	
		Rev	CGCGGATCC -ACCCACGTTGTAAGGTTG	BamHI	
70	fu 961 c-L-...	Fwd	CGCGGATCCCATATG- ATGAAACACTTCCATCC	NdeI	
		Rev	CGCGGATCC -ACCCACGTTGTAAGGTTG	BamHI	
75	fu (961 )- 741(MC58)-His	Fwd	CGCGGATCC -GGAGGGGGTGGTGTGCG	BamHI	
		Rev	CCCGCTCGAG-TTGCTTGGCGGCAAGGC	XhoI	
80	fu (961 )-983-His	Fwd	CGCGGATCC - GGCGGAGGCAGGCACTT	BamHI	
		Rev	CCCGCTCGAG-GAACCGTAGCCTACG	XhoI	
85	fu (961)- Orf46.1- His	Fwd	CGCGGATCCGGTGGTGGTGGT- TCAGATTGGCAAACGATT	BamHI	
		Rev	CCCGCTCGAG-CGTATCATATTCACGTGC	XhoI	
90	fu (961 c-L)- 741(MC58)	Fwd	CGCGGATCC -GGAGGGGGTGGTGTGCG	BamHI	
		Rev	CCCGCTCGAG-TTATTGCTTGGCGGCAAG	XhoI	
95	fu (961c-L )-983	Fwd	CGCGGATCC - GGCGGAGGCAGGCACTT	BamHI	
		Rev	CCCGCTCGAG-TCAGAACCGGTAGCCTAC	XhoI	
100	fu (961c-L)- Orf46.1	Fwd	CGCGGATCCGGTGGTGGTGGT- TCAGATTGGCAAACGATT	BamHI	
		Rev	CCCGCTCGAG-TTACGTATCATATTCACGTGC	XhoI	
105	961-(His/GST)	Fwd	CGCGGATCCCATATG-GCCACAAAGCGACGACG	BamHI-NdeI	

	(MC58)	Rev	<u>CCCGCTCGAG</u> -CCACTCGTAATTGACGCC	Xhol
5	961 Δ1-His	Fwd	CGCGGATCCCATATG-GCCACAAACGACGAC	NdeI
		Rev	<u>CCCGCTCGAG</u> -TGCTTGGCGGAAAGTT	Xhol
10	961a-(His/GST)	Fwd	CGCGGATCCCATATG-GCCACAAACGACGAC	BamHI-NdeI
		Rev	<u>CCCGCTCGAG</u> -TTAGCAATATTATCTTGTTCGTAGC	Xhol
15	961b-(His/GST)	Fwd	CGCGGATCCCATATG-AAAGCAAACCGTGCAGA	BamHI-NdeI
		Rev	<u>CCCGCTCGAG</u> -CCACTCGTAATTGACGCC	Xhol
20	961-His/GST <sup>GATE</sup>	Fwd	ggggacaagtttacaaaaaaggcggctGCAGGCCACAAACGACGACG ATGTTAAAAAACG	<i>attB1</i>
		Rev	ggggaccatgttacaaagaaagtgggTTACCACTCGTAATTGACGC CGACATGGTAGG	<i>attB2</i>
25	982	Fwd	<u>GCGGCCATATG</u> -GCAGCAAAAGACGTACAGTT	NdeI
		Rev	<u>GCGGCCTCGAG</u> -TTACATCATGCCGCCATACCA	Xhol
30	983-His (2996)	Fwd	CGCGGATCCGCTAGC-TTACCGGGCGGGAG	NheI
		Rev	<u>CCCGCTCGAG</u> -GAACCGGTAGCCTACG	Xhol
35	ΔG983-His (2996)	Fwd	CCCTCTAGCTAGC-ACTTCTGCGCCGACTT	NheI
		Rev	<u>CCCGCTCGAG</u> -GAACCGGTAGCCTACG	Xhol
40	983-His	Fwd	CGCGGATCCGCTAGC-TTACCGGGCGGGAG	NheI
		Rev	<u>CCCGCTCGAG</u> -GAACCGGTAGCCTACG	Xhol
45	ΔG983-His	Fwd	CGCGGATCCGCTAGC-ACTTCTGCGCCGACTT	NheI
		Rev	<u>CCCGCTCGAG</u> -GAACCGGTAGCCTACG	Xhol
50	983L	Fwd	CGCGGATCCGCTAGC- CGAACGACCCCAACCTTCCCTACAAAAACTTCAA	NheI
		Rev	<u>CCCGCTCGAG</u> -TCAGAACCGACGTGCCAAGCCGTT	Xhol
55	987-His (MC58)	Fwd	GCCGCCATATGCCCAACTGGAAAGAACGGACG	NdeI
		Rev	GCCGCCTCGAGTAATAAACCTTATGGCAGCAG	Xhol
60	989-(His/GST) (MC58)	Fwd	CGCGGATCCCATATG-TCCGTCCACGCATCCG	BamHI-NdeI
		Rev	<u>CCCGCTCGAG</u> -TTGAATTGTAGGTGTATTG	Xhol
65	989L (MC58)	Fwd	CGCGGATCCCATATG-ACCCCTTCCGCACT	NdeI
		Rev	<u>CCCGCTCGAG</u> -TTATTGAAATTGTAGGTGTAT	Xhol
70	CrgA-His (MC58)	Fwd	CGCGGATCCCATATG-AAAACCAATTAGAAGAA	NdeI
		Rev	<u>CCCGCTCGAG</u> -TCCACAGAGATTGTTCC	Xhol
75	PilC1-ES (MC58)	Fwd	GATGCCGAAGGGCGGG	
		Rev	GCCCAAGCTT-TCAGAACGACTTCACCGC	
80	PilC1-His (MC58)	Fwd	CGCGGATCCCATATG-CAAACCCATAAATACGCTATT	NdeI
		Rev	GCCCAAGCTT-GAAGAACGACTTCACGCCAG	HindIII
85	Δ1PilC1-His (MC58)	Fwd	CGCGGATCCCATATG-GTCTTTCGACAATACCGA	NdeI
		Rev	GCCCAAGCTT-	HindIII
90	PilC1L (MC58)	Fwd	CGCGGATCCCATATG-AATAAAACTTAAAAAGGCGG	NdeI
		Rev	GCCCAAGCTT-TCAGAACGACTTCACCGC	HindIII
95	ΔGTbp2-His (MC58)	Fwd	CGCGAATCCCATATG-TTCGATCTGATTCTGCGA	NdeI
		Rev	<u>CCCGCTCGAG</u> -TCGCACAGGCTGTTGGCG	Xhol
100	Tbp2-His (MC58)	Fwd	CGCGAATCCCATATG-TTGGGCGGAGGCCAG	NdeI
		Rev	<u>CCCGCTCGAG</u> -TCGCACAGGCTGTTGGCG	Xhol
105	Tbp2-His(MC58)	Fwd	CGCGAATCCCATATG-TTGGGCGGAGGCCAG	NdeI
		Rev	<u>CCCGCTCGAG</u> -TCGCACAGGCTGTTGGCG	Xhol

5	<b>NMB0109- (His/GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-GCAAATTGGAGGTGCGC</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAG-TTCGGAGCGGTTGAAGC</u>	Xhol
10	<b>NMB0109L (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-CAACGTCGTATTATAACCC</u>	NdeI
		Rev	<u>CCCGCTCGAG-TTATTCGGAGCGGTTGAAG</u>	Xhol
15	<b>NMB0207- (His/GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-GGCATCAAAGTCGCCATCAACGGCTAC</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAG-TTTGAGCGGGCGCACTTCAAGTCCG</u>	Xhol
20	<b>NMB0462- (His/GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-GGCAGCGAAAAAAC</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAG-GTTGGTGCCGACTTGAT</u>	Xhol
25	<b>NMB0623- (His/GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-GGCAGCGGAAGCGATA</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAG-TTTGCCGCTTGAGCC</u>	Xhol
30	<b>NMB0625 (His- GST)(MC58)</b>	Fwd	<u>CGCGGATCCCATATGGCAAATCCGAAAATACG</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAGCATCCGTACTGTTCG</u>	Xhol
35	<b>NMB0634 (His/GST)(MC58)</b>	Fwd	ggggacaagttgtacaaaaaaaggctCCGACATTACCGTGTACAAC GGCCAACAAAGAA	<i>attB1</i>
		Rev	ggggaccacttgtacaaagaaagctgggCTTATTCATACCGGCTTGCT CAAGCAGCCG	<i>attB2</i>
40	<b>NMB0776- His/GST (MC58) GATE</b>	Fwd	ggggacaagttgtacaaaaaaaggctGATAACGGTTTTCCTGTAA AACGGACAACAA	<i>attB1</i>
		Rev	ggggaccacttgtacaaagaaagctgggCTAGGAAAAATCGTCATCGT TGAAATTGCC	<i>attB2</i>
45	<b>NMB1115- His/GST (MC58) GATE</b>	Fwd	ggggacaagttgtacaaaaaaaggctATGCACCCATCGAAACC	<i>attB1</i>
		Rev	ggggaccacttgtacaaagaaagctgggCTAGTCTGCAGTGCCTC	<i>attB2</i>
50	<b>NMB1343- (His/GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-GGAAATTTCTTATATAGAGGCATTAG</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAG-GTTAATTCTATCAACTCTTAGCAATAAT</u>	Xhol
55	<b>NMB1369 (His- GST (MC58)</b>	Fwd	<u>CGCGGATCCCATATGGCCTGCCAGACGACA</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAGGCCCTCTGCCAAA</u>	Xhol
60	<b>NMB1551 (His- GST)(MC58)</b>	Fwd	<u>CGCGGATCCCATATGGCAGAGATCTGTTGATAA</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAGCGGTTTCCGCCAATG</u>	Xhol
65	<b>NMB1899 (His- GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATGCAGCCGATACGGTC</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAGAACACTTCAACACAAAAT</u>	Xhol
70	<b>NMB2050- (His/GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-TGGTTGCTGATGAAGGGC</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAG-GACTGCTTCATCTTCTGC</u>	Xhol
75	<b>NMB2050L (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-GAACTGATGACTGTTGC</u>	NdeI
		Rev	<u>CCCGCTCGAG-TCAGACTGCTTCATCTTCT</u>	Xhol
80	<b>NMB2159- (His/GST) (MC58)</b>	Fwd	<u>CGCGGATCCCATATG-AGCATTAAAGTAGCGATTAACGGTTCGGC</u>	BamHI-NdeI
		Rev	<u>CCCGCTCGAG-GATTTGCCGCGAAGTATTCAAAGTGC</u>	Xhol
85	<b>fu-ΔG287...-His</b>	Fwd	<u>CGCGGATCCGCTAGC-CCCGATGTTAAATCGGC</u>	NheI

	Rev	<u>CGGGGATCC-ATCCTGCTCTTTTGCCGG</u>	BamHI
5	<b>fu-(ΔG287)-919-His</b>	Fwd <u>CGCGGATCCGGTGGTGGTGGT-</u> CAAAGCAAGAGCATCCAAACC	BamHI
		Rev <u>CCCAAGCTT-TTCGGGCGGTATTGGGCTTC</u>	HindIII
		Fwd <u>CGCGGATCCGGTGGTGGTGGT-</u> GCCACCTACAAAGTGGAC	BamHI
10	<b>fu-(ΔG287)-953-His</b>	Rev <u>GCCCAAGCTT-TTGTGGCTGCCTCGAT</u>	HindIII
		Fwd <u>CGCGGATCCGGTGGTGGTGGT-ACAAGCGACGACG</u>	BamHI
		Rev <u>GCCCAAGCTT-CCACTCGTAATTGACGCC</u>	HindIII
15	<b>fu-(ΔG287)-961-His</b>	Fwd <u>CGCGGATCCGGTGGTGGTGGT-TCAGATTGGCAAACGATT</u>	BamHI
		Rev <u>CCCAAGCTT-CGTATCATATTTCACGTGC</u>	HindIII
		Fwd <u>CCCAAGCTTGGTGGTGGTGGTGGT-</u> TCAGATTGGCAAACGATT	HindIII
20	<b>fu-(ΔG287-919)-Orf46.1-His</b>	Rev <u>CCCGCTCGAG-CGTATCATATTTCACGTGC</u>	Xhol
		Fwd <u>CCCAAGCTTGGTGGTGGTGGTGGT-</u> CAAAGCAAGAGCATCCAAACC	HindIII
		Rev <u>CCCGCTCGAG-CGGGCGGTATTGGGCTT</u>	Xhol
25	<b>fu ΔG287(394.98)-...</b>	Fwd <u>CGCGGATCCGCTAGC-CCCGATGTTAACCGC</u>	NheI
		Rev <u>CGGGGATCC-ATCCTGCTCTTTTGCCGG</u>	BamHI
		Fwd <u>CGCGGATCCGCTAGC-GGACACACTTATTCCGGC</u>	NheI
30	<b>fu Orf1-(Orf46.1)-His</b>	Rev <u>CGCGGATCC-CCAGCGGTAGCCTAATTGAT</u>	
		Fwd <u>CGCGGATCCGGTGGTGGTGGT-</u> TCAGATTGGCAAACGATT	BamHI
		Rev <u>CCCAAGCTT-CGTATCATATTTCACGTGC</u>	HindIII
35	<b>fu (919)-Orf46.1-His</b>	Fwd1 <u>GCGGCGTCGACGGTGGCGGAGGC</u> ACTGGATCCTCAG	Sall
		Fwd2 <u>GGAGGCACTGGATCCTCAGATTGGCAAACGATT</u>	
		Rev <u>CCCGCTCGAG-CGTATCATATTTCACGTGC</u>	Xhol
40	<b>Fu orf46-....</b>	Fwd <u>GGAATTCCATATGTCAGATTGGCAAACGATT</u>	NdeI
		Rev <u>CGCGGATCCCGTATCATATTTCACGTGC</u>	BamHI
		Fwd <u>CGGGGATCCGGGGCGCGGTGGCG</u>	BamHI
45	<b>Fu (orf46)-287-His</b>	Rev <u>CCCAAGCTTATCCTGCTCTTTTGCCGGC</u>	HindIII
		Fwd <u>CGCGGATCCGGTGGTGGTCAAAGCAAGAGCATCCA</u>	BamHI
		Rev <u>CCCAAGCTTGGGCGGTATTGGGCTTC</u>	HindIII
50	<b>Fu (orf46-919)-287-His</b>	Fwd <u>CCCCAAGCTTGGGGCGCGGTGGCG</u>	HindIII
		Rev <u>CCCGCTCGAGATCCTGCTCTTTTGCCGGC</u>	Xhol
		Fwd <u>CCCAAGCTTGGTGGTGGTGGTCAAAGCAAGAGCAT</u>	HindIII
55	<b>Fu (orf46-287)-919-His</b>	Rev <u>CCCGCTCGAGCGGGCGGTATTGGGCTT</u>	Xhol
		Fwd1 <u>GGAGGCACTGGATCCGAGCCACAAACGACGACGA</u>	Xhol
		Fwd2 <u>GCGGCCTCGAG-GGTGGCGGAGGC</u> ACTGGATCCGCAG	
	<b>(ΔG741 )-961c-His</b>	Rev <u>CCCGCTCGAG-ACCCAGCTTGTAAAGTTG</u>	Xhol
		Fwd1 <u>GGAGGCACTGGATCCGAGCCACAAACGACGACGA</u>	Xhol
		Fwd2 <u>GCGGCCTCGAG-GGTGGCGGAGGC</u> ACTGGATCCGCAG	
	<b>(ΔG741 )-961-His</b>	Rev <u>CCCGCTCGAG-CCACTCGTAATTGACGCC</u>	Xhol

5	<b>(ΔG741 )-983-His</b>	Fwd	<b>GCGGCCTCGAG-</b> GGATCCGGCGGAGGCAGCACTTCTGCG	XbaI
		Rev	<b>CCCGCTCGAG-</b> GAACCGGTAGCCTACG	XbaI
10	<b>(ΔG741 )-orf46.1-His</b>	Fwd1	GGAGGCACTGGATCCTCAGATTGGCAAACGATTG	Sall
		Fwd2	<b>GCGGCCTCGAG-</b> CGGTGGCGGAGGCAGTGATCCTCAGA	
		Rev	<b>CCCGCTCGAG-</b> CGTATCATATTCACGTGC	XbaI
15	<b>(ΔG983)-741(MC58) -His</b>	Fwd	<b>GCGGCCTCGAG-</b> GGATCCGGAGGGGGTGGTGTGCC	XbaI
		Rev	<b>CCCGCTCGAG-</b> TTGCTTGGCGGCAAG	XbaI
20	<b>(ΔG983)-961c-His</b>	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XbaI
		Fwd2	<b>GCGGCCTCGAG-</b> GGTGGCGGAGGCAGTGATCCGAG	
		Rev	<b>CCCGCTCGAG-</b> ACCCAGCTTGTAAAGTTG	XbaI
25	<b>(ΔG983)-961-His</b>	Fwd1	GGAGGCACTGGATCCGCAGCCACAAACGACGACGA	XbaI
		Fwd2	<b>GCGGCCTCGAG-</b> GGTGGCGGAGGCAGTGATCCGAG	
		Rev	<b>CCCGCTCGAG-</b> CCACTCGTAATTGACGCC	XbaI
30	<b>(ΔG983)- Orf46.1-His</b>	Fwd1	GGAGGCACTGGATCCTCAGATTGGCAAACGATTG	Sall
		Fwd2	<b>GCGGCCTCGAG-</b> CGGTGGCGGAGGCAGTGATCCTCAGA	
		Rev	<b>CCCGCTCGAG-</b> CGTATCATATTCACGTGC	XbaI

\* This primer was used as a Reverse primer for all the C terminal fusions of 287 to the His-tag.

§ Forward primers used in combination with the 287-His Reverse primer.

NB – All PCR reactions use strain 2996 unless otherwise specified (e.g. strain MC58)

[0293] In all constructs starting with an ATG not followed by a unique *Nhe*I site, the ATG codon is part of the *Nde*I site used for cloning. The constructs made using *Nhe*I as a cloning site at the 5' end (e.g. all those containing 287 at the N-terminus) have two additional codons (GCT AGC) fused to the coding sequence of the antigen.

#### Preparation of chromosomal DNA templates

[0294] *N.meningitidis* strains 2996, MC58, 394.98, 1000 and BZ232 (and others) were grown to exponential phase in 100ml of GC medium, harvested by centrifugation, and resuspended in 5ml buffer (20% w/v sucrose, 50mM Tris-HCl, 50mM EDTA, pH8). After 10 minutes incubation on ice, the bacteria were lysed by adding 10ml of lysis solution (50mM NaCl, 1% Na-Sarkosyl, 50µg/ml Proteinase K), and the suspension incubated at 37°C for 2 hours. Two phenol extractions (equilibrated to pH 8) and one CHCl<sub>3</sub>/isoamylalcohol (24:1) extraction were performed. DNA was precipitated by addition of 0.3M sodium acetate and 2 volumes of ethanol, and collected by centrifugation. The pellet was washed once with 70%(v/v) ethanol and redissolved in 4.0ml TE buffer (10mM Tris-HCl, 1mM EDTA, pH 8.0). The DNA concentration was measured by reading OD<sub>260</sub>.

#### PCR Amplification

[0295] The standard PCR protocol was as follows: 200ng of genomic DNA from 2996, MC581000, or BZ232 strains or 10ng of plasmid DNA preparation of recombinant clones were used as template in the presence of 40µM of each oligonucleotide primer, 400-800 µM dNTPs solution, 1x PCR buffer (including 1.5mM MgCl<sub>2</sub>), 2.5 units *Taq*/DNA polymerase (using Perkin-Elmer AmpliTaq, Boerhingher Mannheim Expand™ Long Template).

[0296] After a preliminary 3 minute incubation of the whole mix at 95°C, each sample underwent a two-step amplification: the first 5 cycles were performed using the hybridisation temperature that excluded the restriction enzyme tail of the primer (*T*<sub>m1</sub>). This was followed by 30 cycles according to the hybridisation temperature calculated for the whole length oligos (*T*<sub>m2</sub>). Elongation times, performed at 68°C or 72°C, varied according to the length of the Orf to be amplified. In the case of Orf1 the elongation time, starting from 3 minutes, was increased by 15 seconds each cycle. The cycles were completed with a 10 minute extension step at 72°C.

[0297] The amplified DNA was either loaded directly on a 1% agarose gel. The DNA fragment corresponding to the band of correct size was purified from the gel using the Qiagen Gel Extraction Kit, following the manufacturer's protocol.

**Digestion of PCR fragments and of the cloning vectors**

[0298] The purified DNA corresponding to the amplified fragment was digested with the appropriate restriction enzymes for cloning into pET-21b+, pET22b+ or pET-24b+. Digested fragments were purified using the QIAquick PCR purification kit (following the manufacturer's instructions) and eluted with either H<sub>2</sub>O or 10mM Tris, pH 8.5. Plasmid vectors were digested with the appropriate restriction enzymes, loaded onto a 1.0% agarose gel and the band corresponding to the digested vector purified using the Qiagen QIAquick Gel Extraction Kit.

**Cloning**

[0299] The fragments corresponding to each gene, previously digested and purified, were ligated into pET21 b+, pET22b+ or pET-24b+. A molar ratio of 3:1 fragment/vector was used with T4 DNA ligase in the ligation buffer supplied by the manufacturer.

[0300] Recombinant plasmid was transformed into competent *E.coli* DH5 or HB101 by incubating the ligation reaction solution and bacteria for 40 minutes on ice, then at 37°C for 3 minutes.

[0301] This was followed by the addition of 800μl LB broth and incubation at 37°C for 20 minutes. The cells were centrifuged at maximum speed in an Eppendorf microfuge, resuspended in approximately 200μl of the supernatant and plated onto LB ampicillin (100mg/ml) agar.

[0302] Screening for recombinant clones was performed by growing randomly selected colonies overnight at 37°C in 4.0ml of LB broth + 100μg/ml ampicillin. Cells were pelleted and plasmid DNA extracted using the Qiagen QIAprep Spin Miniprep Kit, following the manufacturer's instructions. Approximately 1μg of each individual miniprep was digested with the appropriate restriction enzymes and the digest loaded onto a 1-1.5% agarose gel (depending on the expected insert size), in parallel with the molecular weight marker (1kb DNA Ladder, GIBCO). Positive clones were selected on the basis of the size of insert.

**Expression**

[0303] After cloning each gene into the expression vector, recombinant plasmids were transformed into *E.coli* strains suitable for expression of the recombinant protein. 1μl of each construct was used to transform *E.coli* BL21-DE3 as described above. Single recombinant colonies were inoculated into 2ml LB+Amp (100μg/ml), incubated at 37°C overnight, then diluted 1:30 in 20ml of LB+Amp (100μg/ml) in 100ml flasks, to give an OD<sub>600</sub> between 0.1 and 0.2. The flasks were incubated at 30°C or at 37°C in a gyratory water bath shaker until OD<sub>600</sub> indicated exponential growth suitable for induction of expression (0.4-0.8 OD). Protein expression was induced by addition of 1.0mM IPTG. After 3 hours incubation at 30°C or 37°C the OD<sub>600</sub> was measured and expression examined. 1.0ml of each sample was centrifuged in a microfuge, the pellet resuspended in PBS and analysed by SDS-PAGE and Coomassie Blue staining.

**Gateway cloning and expression**

[0304] Sequences labelled GATE were cloned and expressed using the GATEWAY Cloning Technology (GIBCO-BRL). Recombinational cloning (RC) is based on the recombination reactions that mediate the integration and excision of phage into and from the *E.coli* genome, respectively. The integration involves recombination of the *attP* site of the phage DNA within the *attB* site located in the bacterial genome (BP reaction) and generates an integrated phage genome flanked by *attL* and *attR* sites. The excision recombines *attL* and *attR* sites back to *attP* and *attB* sites (LR reaction). The integration reaction requires two enzymes [the phage protein Integrase (Int) and the bacterial protein integration host factor (IHF)] (BP clonase). The excision reaction requires Int, IHF, and an additional phage enzyme, Excisionase (Xis) (LR clonase). Artificial derivatives of the 25-bp bacterial *attB* recombination site, referred to as B 1 and B2, were added to the 5' end of the primers used in PCR reactions to amplify Neisserial ORFs. The resulting products were BP cloned into a "Donor vector" containing complementary derivatives of the phage *attP* recombination site (P1 and P2) using BP clonase. The resulting "Entry clones" contain ORFs flanked by derivatives of the *attL* site (L1 and L2) and were subcloned into expression "destination vectors" which contain derivatives of the *attL*-compatible *attR* sites (R1 and R2) using LR clonase. This resulted in "expression clones" in which ORFs are flanked by B1 and B2 and fused in frame to the GST or His N terminal tags.

[0305] The *E. coli* strain used for GATEWAY expression is BL21-SI. Cells of this strain are induced for expression of the T7 RNA polymerase by growth in medium containing salt (0.3 M NaCl).

[0306] Note that this system gives N-terminus His tags.

***Preparation of membrane proteins.***

[0307] Fractions composed principally of either inner, outer or total membrane were isolated in order to obtain recombinant proteins expressed with membrane-localisation leader sequences. The method for preparation of membrane fractions, enriched for recombinant proteins, was adapted from Filip et. al. [J.Bact. (1973) 115:717-722] and Davies et. al. [J. Immunol. Meth. (1990) 143:215-225]. Single colonies harbouring the plasmid of interest were grown overnight at 37°C in 20 ml of LB/Amp (100 µg/ml) liquid culture. Bacteria were diluted 1:30 in 1.0 L of fresh medium and grown at either 30°C or 37°C until the OD<sub>550</sub> reached 0.6-0.8. Expression of recombinant protein was induced with IPTG at a final concentration of 1.0 mM. After incubation for 3 hours, bacteria were harvested by centrifugation at 8000g for 15 minutes at 4°C and resuspended in 20 ml of 20 mM Tris-HCl (pH 7.5) and complete protease inhibitors (Boehringer-Mannheim). All subsequent procedures were performed at 4°C or on ice.

[0308] Cells were disrupted by sonication using a Branson Sonifier 450 and centrifuged at 5000g for 20 min to sediment unbroken cells and inclusion bodies. The supernatant, containing membranes and cellular debris, was centrifuged at 50000g (Beckman Ti50, 29000rpm) for 75 min, washed with 20 mM Bis-tris propane (pH 6.5), 1.0 M NaCl, 10% (v/v) glycerol and sedimented again at 50000g for 75 minutes. The pellet was resuspended in 20mM Tris-HCl (pH 7.5), 2.0% (v/v) Sarkosyl, complete protease inhibitor (1.0 mM EDTA, final concentration) and incubated for 20 minutes to dissolve inner membrane. Cellular debris was pelleted by centrifugation at 5000g for 10 min and the supernatant centrifuged at 75000g for 75 minutes (Beckman Ti50, 33000rpm). Proteins 008L and 519L were found in the supernatant suggesting inner membrane localisation. For these proteins both inner and total membrane fractions (washed with NaCl as above) were used to immunise mice. Outer membrane vesicles obtained from the 75000g pellet were washed with 20 mM Tris-HCl (pH 7.5) and centrifuged at 75000g for 75 minutes or overnight. The OMV was finally resuspended in 500 µl of 20 mM Tris-HCl (pH 7.5), 10% v/v glycerol. Orf1L and Orf40L were both localised and enriched in the outer membrane fraction which was used to immunise mice. Protein concentration was estimated by standard Bradford Assay (Bio-Rad), while protein concentration of inner membrane fraction was determined with the DC protein assay (Bio-Rad). Various fractions from the isolation procedure were assayed by SDS-PAGE.

***Purification of His-tagged proteins***

[0309] Various forms of 287 were cloned from strains 2996 and MC58. They were constructed with a C-terminus His-tagged fusion and included a mature form (aa 18-427), constructs with deletions ( $\Delta$ 1,  $\Delta$  2,  $\Delta$ 3 and  $\Delta$ 4) and clones composed of either B or C domains. For each clone purified as a His-fusion, a single colony was streaked and grown overnight at 37°C on a LB/Amp (100 µg/ml) agar plate. An isolated colony from this plate was inoculated into 20ml of LB/Amp (100 µg/ml) liquid medium and grown overnight at 37°C with shaking. The overnight culture was diluted 1:30 into 1.0 L LB/Amp (100 µg/ml) liquid medium and allowed to grow at the optimal temperature (30 or 37°C) until the OD<sub>550</sub> reached 0.6-0.8. Expression of recombinant protein was induced by addition of IPTG (final concentration 1.0mM) and the culture incubated for a further 3 hours. Bacteria were harvested by centrifugation at 8000g for 15 min at 4°C. The bacterial pellet was resuspended in 7.5 ml of either (i) cold buffer A (300 mM NaCl, 50 mM phosphate buffer, 10 mM imidazole, pH 8.0) for soluble proteins or (ii) buffer B (10mM Tris-HCl, 100 mM phosphate buffer, pH 8.8 and, optionally, 8M urea) for insoluble proteins. Proteins purified in a soluble form included 287-His,  $\Delta$ 1,  $\Delta$ 2,  $\Delta$ 3 and  $\Delta$ 4287-His,  $\Delta$ 4287MC58-His, 287c-His and 287cMC58-His. Protein 287bMC58-His was insoluble and purified accordingly. Cells were disrupted by sonication on ice four times for 30 sec at 40W using a Branson sonifier 450 and centrifuged at 13000xg for 30 min at 4°C. For insoluble proteins, pellets were resuspended in 2.0 ml buffer C (6 M guanidine hydrochloride, 100 mM phosphate buffer, 10 mM Tris- HCl, pH 7.5 and treated with 10 passes of a Dounce homogenizer. The homogenate was centrifuged at 13000g for 30 min and the supernatant retained. Supernatants for both soluble and insoluble preparations were mixed with 150µl Ni<sup>2+</sup>-resin (previously equilibrated with either buffer A or buffer B, as appropriate) and incubated at room temperature with gentle agitation for 30 min. The resin was Chelating Sepharose Fast Flow (Pharmacia), prepared according to the manufacturer's protocol. The batch-wise preparation was centrifuged at 700g for 5 min at 4°C and the supernatant discarded. The resin was washed twice (batch-wise) with 10ml buffer A or B for 10 min, resuspended in 1.0 ml buffer A or B and loaded onto a disposable column. The resin continued to be washed with either (i) buffer A at 4°C or (ii) buffer B at room temperature, until the OD<sub>280</sub> of the flow-through reached 0.02-0.01. The resin was further washed with either (i) cold buffer C (300mM NaCl, 50mM phosphate buffer, 20mM imidazole, pH 8.0) or (ii) buffer D (10mM Tris-HCl, 100mM phosphate buffer, pH 6.3 and, optionally, 8M urea) until OD<sub>280</sub> of the flow-through reached 0.02-0.01. The His-fusion protein was eluted by addition of 700µl of either (i) cold elution buffer A (300 mM NaCl, 50mM phosphate buffer, 250 mM imidazole, pH 8.0) or (ii) elution buffer B (10 mM Tris-HCl, 100 mM phosphate buffer, pH 4.5 and, optionally, 8M urea) and fractions collected until the OD<sub>280</sub> indicated all the recombinant protein was obtained. 20µl aliquots of each elution fraction were analysed by SDS-PAGE. Protein concentrations were estimated using the Bradford assay.

***Renaturation of denatured His-fusion proteins.***

[0310] Denaturation was required to solubilize 287bMC8, so a renaturation step was employed prior to immunisation. Glycerol was added to the denatured fractions obtained above to give a final concentration of 10% v/v. The proteins were diluted to 200 µg/ml using dialysis buffer I (10% v/v glycerol, 0.5M arginine, 50 mM phosphate buffer, 5.0 mM reduced glutathione, 0.5 mM oxidised glutathione, 2.0M urea, pH 8.8) and dialysed against the same buffer for 12-14 hours at 4°C. Further dialysis was performed with buffer II (10% v/v glycerol, 0.5M arginine, 50mM phosphate buffer, 5.0mM reduced glutathione, 0.5mM oxidised glutathione, pH 8.8) for 12-14 hours at 4°C. Protein concentration was estimated using the formula:

10

$$\text{Protein (mg/ml)} = (1.55 \times OD_{280}) - (0.76 \times OD_{260})$$

15 ***Amino acid sequence analysis.***

[0311] Automated sequence analysis of the NH<sub>2</sub>-terminus of proteins was performed on a Beckman sequencer (LF 3000) equipped with an on-line phenylthiohydantoin-amino acid analyser (System Gold) according to the manufacturer's recommendations.

20

***Immunization***

[0312] Balb/C mice were immunized with antigens on days 0, 21 and 35 and sera analyzed at day 49.

25

***Sera analysis - ELISA***

[0313] The acapsulated MenB M7 and the capsulated strains were plated on chocolate agar plates and incubated overnight at 37°C with 5% CO<sub>2</sub>. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.4-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and bacteria were washed twice with PBS, resuspended in PBS containing 0.025% formaldehyde, and incubated for 1 hour at 37°C and then overnight at 4°C with stirring. 100µl bacterial cells were added to each well of a 96 well Greiner plate and incubated overnight at 4°C. The wells were then washed three times with PBT washing buffer (0.1% Tween-20 in PBS). 200µl of saturation buffer (2.7% polyvinylpyrrolidone 10 in water) was added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT. 200µl of diluted sera (Dilution buffer: 1% BSA, 0.1 % Tween-20, 0.1 % NaN<sub>3</sub> in PBS) were added to each well and the plates incubated for 2 hours at 37°C. Wells were washed three times with PBT buffer. 100µl of HRP-conjugated rabbit anti-mouse (Dako) serum diluted 1:2000 in dilution buffer were added to each well and the plates were incubated for 90 minutes at 37°C. Wells were washed three times with PBT buffer. 100µl of substrate buffer for HRP (25ml of citrate buffer pH5, 10mg of O-phenyldiamine and 10µl of H<sub>2</sub>O<sub>2</sub>) were added to each well and the plates were left at room temperature for 20 minutes. 100µl 12.5% H<sub>2</sub>SO<sub>4</sub> was added to each well and OD<sub>490</sub> was followed. The ELISA titers were calculated arbitrarily as the dilution of sera which gave an OD<sub>490</sub> value of 0.4 above the level of preimmune sera. The ELISA was considered positive when the dilution of sera with OD<sub>490</sub> of 0.4 was higher than 1:400.

45

***Sera analysis - FACS Scan bacteria binding assay***

[0314] The acapsulated MenB M7 strain was plated on chocolate agar plates and incubated overnight at 37°C with 5% CO<sub>2</sub>. Bacterial colonies were collected from the agar plates using a sterile dracon swab and inoculated into 4 tubes containing 8ml each Mueller-Hinton Broth (Difco) containing 0.25% glucose. Bacterial growth was monitored every 30 minutes by following OD<sub>620</sub>. The bacteria were let to grow until the OD reached the value of 0.35-0.5. The culture was centrifuged for 10 minutes at 4000rpm. The supernatant was discarded and the pellet was resuspended in blocking buffer (1% BSA in PBS, 0.4% NaN<sub>3</sub>) and centrifuged for 5 minutes at 4000rpm. Cells were resuspended in blocking buffer to reach OD<sub>620</sub> of 0.05. 100µl bacterial cells were added to each well of a Costar 96 well plate. 100µl of diluted (1:100, 1:200, 1:400) sera (in blocking buffer) were added to each well and plates incubated for 2 hours at 4°C. Cells were centrifuged for 5 minutes at 4000rpm, the supernatant aspirated and cells washed by addition of 200µl/well of blocking buffer in each well. 100µl of R-Phicoerytrin conjugated F(ab)<sub>2</sub> goat anti-mouse, diluted 1:100, was added to each well and plates incubated for 1 hour at 4°C. Cells were spun down by centrifugation at 4000rpm for 5 minutes and washed by addition of 200µl/well of blocking buffer. The supernatant was aspirated and cells resuspended in 200µl/

well of PBS, 0.25% formaldehyde. Samples were transferred to FACSscan tubes and read. The condition for FACSscan (Laser Power 15mW) setting were: FL2 on; FSC-H threshold:92; FSC PMT Voltage: E 01; SSC PMT: 474; Amp. Gains 6.1; FL-2 PMT: 586; compensation values: 0.

5      ***Sera analysis - bactericidal assay***

[0315] *N. meningitidis* strain 2996 was grown overnight at 37°C on chocolate agar plates (starting from a frozen stock) with 5% CO<sub>2</sub>. Colonies were collected and used to inoculate 7ml Mueller-Hinton broth, containing 0.25% glucose to reach an OD<sub>620</sub> of 0.05-0.08. The culture was incubated for approximately 1.5 hours at 37 degrees with shaking until the OD<sub>620</sub> reached the value of 0.23-0.24. Bacteria were diluted in 50mM Phosphate buffer pH 7.2 containing 10mM MgCl<sub>2</sub>, 10mM CaCl<sub>2</sub> and 0.5% (w/v) BSA (assay buffer) at the working dilution of 10<sup>5</sup> CFU/ml. The total volume of the final reaction mixture was 50 µl with 25 µl of serial two fold dilution of test serum, 12.5 µl of bacteria at the working dilution, 12.5 µl of baby rabbit complement (final concentration 25%).

[0316] Controls included bacteria incubated with complement serum, immune sera incubated with bacteria and with complement inactivated by heating at 56°C for 30'. Immediately after the addition of the baby rabbit complement, 10µl of the controls were plated on Mueller-Hinton agar plates using the tilt method (time 0). The 96-wells plate was incubated for 1 hour at 37°C with rotation. 7µl of each sample were plated on Mueller-Hinton agar plates as spots, whereas 10µl of the controls were plated on Mueller-Hinton agar plates using the tilt method (time 1). Agar plates were incubated for 18 hours at 37 degrees and the colonies corresponding to time 0 and time 1 were counted.

20     ***Sera analysis - western blots***

[0317] Purified proteins (500ng/lane), outer membrane vesicles (5µg) and total cell extracts (25µg) derived from MenB strain 2996 were loaded onto a 12% SDS-polyacrylamide gel and transferred to a nitrocellulose membrane. The transfer was performed for 2 hours at 150mA at 4°C, using transfer buffer (0.3% Tris base, 1.44% glycine, 20% (v/v) methanol). The membrane was saturated by overnight incubation at 4°C in saturation buffer (10% skimmed milk, 0.1% Triton X100 in PBS). The membrane was washed twice with washing buffer (3% skimmed milk, 0.1% Triton X100 in PBS) and incubated for 2 hours at 37°C with mice sera diluted 1:200 in washing buffer. The membrane was washed twice and incubated for 90 minutes with a 1:2000 dilution of horseradish peroxidase labelled anti-mouse Ig. The membrane was washed twice with 0.1% Triton X100 in PBS and developed with the Opti-4CN Substrate Kit (Bio-Rad). The reaction was stopped by adding water.

[0318] The OMVs were prepared as follows: *N. meningitidis* strain 2996 was grown overnight at 37 degrees with 5% CO<sub>2</sub> on 5 GC plates, harvested with a loop and resuspended in 10 ml of 20mM Tris-HCl pH 7.5, 2 mM EDTA. Heat inactivation was performed at 56°C for 45 minutes and the bacteria disrupted by sonication for 5 minutes on ice (50% duty cycle, 50% output, Branson sonifier 3 mm microtip). Unbroken cells were removed by centrifugation at 5000g for 10 minutes, the supernatant containing the total cell envelope fraction recovered and further centrifuged overnight at 50000g at the temperature of 4°C . The pellet containing the membranes was resuspended in 2% sarkosyl, 20mM Tris-HCl pH 7.5, 2 mM EDTA and incubated at room temperature for 20 minutes to solubilise the inner membranes. The suspension was centrifuged at 10000g for 10 minutes to remove aggregates, the supernatant was further centrifuged at 50000g for 3 hours. The pellet, containing the outer membranes was washed in PBS and resuspended in the same buffer. Protein concentration was measured by the D.C. Bio-Rad Protein assay (Modified Lowry method), using BSA as a standard.

[0319] Total cell extracts were prepared as follows: *N. meningitidis* strain 2996 was grown overnight on a GC plate, harvested with a loop and resuspended in 1ml of 20mM Tris-HCl. Heat inactivation was performed at 56°C for 30 minutes.

45     ***961 domain studies***

[0320] Cellular fractions preparation Total lysate, periplasm, supernatant and OMV of *E.coli*clones expressing different domains of 961 were prepared using bacteria from over-night cultures or after 3 hours induction with IPTG. Briefly, the periplasm were obtained suspending bacteria in saccarose 25% and Tris 50mM (pH 8) with polimixine 100µg/ml. After 1hr at room temperature bacteria were centrifuged at 13000rpm for 15 min and the supernatant were collected. The culture supernatant were filtered with 0.2µm and precipitated with TCA 50% in ice for two hours. After centrifugation (30 min at 13000 rp) pellets were rinsed twice with ethanol 70% and suspended in PBS. The OMV preparation was performed as previously described. Each cellular fraction were analyzed in SDS-PAGE or in Western Blot using the polyclonal anti-serum raised against GST-961.

[0321] Adhesion assay Chang epithelial cells (Wong-Kilbourne derivative, clone 1-5c-4, human conjunctiva) were maintained in DMEM (Gibco) supplemented with 10% heat-inactivated FCS, 15mM L-glutamine and antibiotics.

[0322] For the adherence assay, sub-confluent culture of Chang epithelial cells were rinsed with PBS and treated with

trypsin-EDTA (Gibco), to release them from the plastic support. The cells were then suspended in PBS, counted and dilute in PBS to  $5 \times 10^5$  cells/ml.

**[0323]** Bacteria from over-night cultures or after induction with IPTG, were pelleted and washed twice with PBS by centrifuging at 13000 for 5 min. Approximately  $2-3 \times 10^8$  (cfu) were incubated with 0.5 mg/ml FITC (Sigma) in 1ml buffer containing 50mM NaHCO<sub>3</sub> and 100mM NaCl pH 8, for 30 min at room temperature in the dark. FITC-labeled bacteria were wash 2-3 times and suspended in PBS at  $1-1.5 \times 10^9$ /ml. 200µl of this suspension ( $2-3 \times 10^8$ ) were incubated with 200µl ( $1 \times 10^5$ ) epithelial cells for 30min a 37°C. Cells were than centrifuged at 2000rpm for 5 min to remove non-adherent bacteria, suspended in 200µl of PBS, transferred to FACScan tubes and read

10 Annex to the application documents - subsequently filed sequences listing

**[0324]**

15

20

25

30

35

40

45

50

55

## SEQUENCE LISTING

&lt;110&gt; Chiron SRL

5 &lt;120&gt; Heterologous Expression of Neisserial Proteins

&lt;130&gt; P044747EP

&lt;140&gt; 06076718.3

&lt;141&gt; 2001-02-28

10 &lt;150&gt; 0004695.3

&lt;151&gt; 2000-02-28

&lt;150&gt; 0027675.8

&lt;151&gt; 2000-11-13

15 &lt;160&gt; 620

&lt;170&gt; Seqwin99, version 1.02

&lt;210&gt; 1

20 &lt;211&gt; 441

&lt;212&gt; PRT

&lt;213&gt; Neisseria meningitidis

&lt;400&gt; 1

Met Lys Lys Tyr Leu Phe Arg Ala Ala Leu Tyr Gly Ile Ala Ala Ala  
1 5 10 15Ile Leu Ala Ala Cys Gln Ser Lys Ser Ile Gln Thr Phe Pro Gln Pro  
20 25 30Asp Thr Ser Val Ile Asn Gly Pro Asp Arg Pro Val Gly Ile Pro Asp  
35 40 45Pro Ala Gly Thr Thr Val Gly Gly Gly Ala Val Tyr Thr Val Val  
50 55 60Pro His Leu Ser Leu Pro His Trp Ala Ala Gln Asp Phe Ala Lys Ser  
65 70 75 80Leu Gln Ser Phe Arg Leu Gly Cys Ala Asn Leu Lys Asn Arg Gln Gly  
85 90 95Trp Gln Asp Val Cys Ala Gln Ala Phe Gln Thr Pro Val His Ser Phe  
100 105 11040 Gln Ala Lys Gln Phe Phe Glu Arg Tyr Phe Thr Pro Trp Gln Val Ala  
115 120 125Gly Asn Gly Ser Leu Ala Gly Thr Val Thr Gly Tyr Tyr Glu Pro Val  
130 135 140Leu Lys Gly Asp Asp Arg Arg Thr Ala Gln Ala Arg Phe Pro Ile Tyr  
145 150 155 160Gly Ile Pro Asp Asp Phe Ile Ser Val Pro Leu Pro Ala Gly Leu Arg  
165 170 17550 Ser Gly Lys Ala Leu Val Arg Ile Arg Gln Thr Gly Lys Asn Ser Gly  
180 185 190Thr Ile Asp Asn Thr Gly Gly Thr His Thr Ala Asp Leu Ser Arg Phe  
195 200 205Pro Ile Thr Ala Arg Thr Thr Ala Ile Lys Gly Arg Phe Glu Gly Ser  
210 215 220

## EP 1 790 660 A2

Arg Phe Leu Pro Tyr His Thr Arg Asn Gln Ile Asn Gly Gly Ala Leu  
 225 230 235 240  
 5 Asp Gly Lys Ala Pro Ile Leu Gly Tyr Ala Glu Asp Pro Val Glu Leu  
 245 250 255  
 Phe Phe Met His Ile Gln Gly Ser Gly Arg Leu Lys Thr Pro Ser Gly  
 260 265 270  
 10 Lys Tyr Ile Arg Ile Gly Tyr Ala Asp Lys Asn Glu His Pro Tyr Val  
 275 280 285  
 Ser Ile Gly Arg Tyr Met Ala Asp Lys Gly Tyr Leu Lys Leu Gly Gln  
 290 295 300  
 15 Thr Ser Met Gln Gly Ile Lys Ala Tyr Met Arg Gln Asn Pro Gln Arg  
 305 310 315 320  
 Leu Ala Glu Val Leu Gly Gln Asn Pro Ser Tyr Ile Phe Phe Arg Glu  
 325 330 335  
 20 Leu Ala Gly Ser Ser Asn Asp Gly Pro Val Gly Ala Leu Gly Thr Pro  
 340 345 350  
 Leu Met Gly Glu Tyr Ala Gly Ala Val Asp Arg His Tyr Ile Thr Leu  
 355 360 365  
 25 Gly Ala Pro Leu Phe Val Ala Thr Ala His Pro Val Thr Arg Lys Ala  
 370 375 380  
 Leu Asn Arg Leu Ile Met Ala Gln Asp Thr Gly Ser Ala Ile Lys Gly  
 385 390 395 400  
 Ala Val Arg Val Asp Tyr Phe Trp Gly Tyr Gly Asp Glu Ala Gly Glu  
 405 410 415  
 30 Leu Ala Gly Lys Gln Lys Thr Thr Gly Tyr Val Trp Gln Leu Leu Pro  
 420 425 430  
 Asn Gly Met Lys Pro Glu Tyr Arg Pro  
 435 440  
 35 <210> 2  
 <211> 420  
 <212> PRT  
 <213> Neisseria meningitidis  
 40 <400> 2  
 Gln Ser Lys Ser Ile Gln Thr Phe Pro Gln Pro Asp Thr Ser Val Ile  
 1 5 10 15  
 Asn Gly Pro Asp Arg Pro Val Gly Ile Pro Asp Pro Ala Gly Thr Thr  
 20 25 30  
 45 Val Gly Gly Gly Ala Val Tyr Thr Val Val Pro His Leu Ser Leu  
 35 40 45  
 Pro His Trp Ala Ala Gln Asp Phe Ala Lys Ser Leu Gln Ser Phe Arg  
 50 55 60  
 50 Leu Gly Cys Ala Asn Leu Lys Asn Arg Gln Gly Trp Gln Asp Val Cys  
 65 70 75 80  
 Ala Gln Ala Phe Gln Thr Pro Val His Ser Phe Gln Ala Lys Gln Phe  
 85 90 95  
 55 Phe Glu Arg Tyr Phe Thr Pro Trp Gln Val Ala Gly Asn Gly Ser Leu  
 100 105 110

## EP 1 790 660 A2

Ala Gly Thr Val Thr Gly Tyr Tyr Glu Pro Val Leu Lys Gly Asp Asp  
 115 120 125  
 5 Arg Arg Thr Ala Gln Ala Arg Phe Pro Ile Tyr Gly Ile Pro Asp Asp  
 130 135 140  
 Phe Ile Ser Val Pro Leu Pro Ala Gly Leu Arg Ser Gly Lys Ala Leu  
 145 150 155 160  
 10 Val Arg Ile Arg Gln Thr Gly Lys Asn Ser Gly Thr Ile Asp Asn Thr  
 165 170 175  
 Gly Gly Thr His Thr Ala Asp Leu Ser Arg Phe Pro Ile Thr Ala Arg  
 180 185 190  
 15 Thr Thr Ala Ile Lys Gly Arg Phe Glu Gly Ser Arg Phe Leu Pro Tyr  
 195 200 205  
 His Thr Arg Asn Gln Ile Asn Gly Gly Ala Leu Asp Gly Lys Ala Pro  
 210 215 220  
 20 Ile Leu Gly Tyr Ala Glu Asp Pro Val Glu Leu Phe Phe Met His Ile  
 225 230 235 240  
 Gln Gly Ser Gly Arg Leu Lys Thr Pro Ser Gly Lys Tyr Ile Arg Ile  
 245 250 255  
 25 Gly Tyr Ala Asp Lys Asn Glu His Pro Tyr Val Ser Ile Gly Arg Tyr  
 260 265 270  
 Met Ala Asp Lys Gly Tyr Leu Lys Leu Gly Gln Thr Ser Met Gln Gly  
 275 280 285  
 30 Ile Lys Ala Tyr Met Arg Gln Asn Pro Gln Arg Leu Ala Glu Val Leu  
 290 295 300  
 Gly Gln Asn Pro Ser Tyr Ile Phe Phe Arg Glu Leu Ala Gly Ser Ser  
 305 310 315 320  
 Asn Asp Gly Pro Val Gly Ala Leu Gly Thr Pro Leu Met Gly Glu Tyr  
 325 330 335  
 35 Ala Gly Ala Val Asp Arg His Tyr Ile Thr Leu Gly Ala Pro Leu Phe  
 340 345 350  
 Val Ala Thr Ala His Pro Val Thr Arg Lys Ala Leu Asn Arg Leu Ile  
 355 360 365  
 40 Met Ala Gln Asp Thr Gly Ser Ala Ile Lys Gly Ala Val Arg Val Asp  
 370 375 380  
 Tyr Phe Trp Gly Tyr Gly Asp Glu Ala Gly Glu Leu Ala Gly Lys Gln  
 385 390 395 400  
 45 Lys Thr Thr Gly Tyr Val Trp Gln Leu Leu Pro Asn Gly Met Lys Pro  
 405 410 415  
 Glu Tyr Arg Pro  
 420  
 50 <210> 3  
 <211> 440  
 <212> PRT  
 <213> Artificial Sequence  
 55 <220>  
 <223> 919  
 <400> 3

Met Lys Thr Phe Phe Lys Thr Leu Ser Ala Ala Ala Leu Ala Leu Ile  
 1 5 10 15  
 5 Leu Ala Ala Cys Gln Ser Lys Ser Ile Gln Thr Phe Pro Gln Pro Asp  
 20 25 30  
 Thr Ser Val Ile Asn Gly Pro Asp Arg Pro Val Gly Ile Pro Asp Pro  
 35 40 45  
 10 Ala Gly Thr Thr Val Gly Gly Gly Ala Val Tyr Thr Val Val Pro  
 50 55 60  
 His Leu Ser Leu Pro His Trp Ala Ala Gln Asp Phe Ala Lys Ser Leu  
 65 70 75 80  
 15 Gln Ser Phe Arg Leu Gly Cys Ala Asn Leu Lys Asn Arg Gln Gly Trp  
 85 90 95  
 Gln Asp Val Cys Ala Gln Ala Phe Gln Thr Pro Val His Ser Phe Gln  
 100 105 110  
 20 Ala Lys Gln Phe Phe Glu Arg Tyr Phe Thr Pro Trp Gln Val Ala Gly  
 115 120 125  
 Asn Gly Ser Leu Ala Gly Thr Val Thr Gly Tyr Tyr Glu Pro Val Leu  
 130 135 140  
 25 Lys Gly Asp Asp Arg Arg Thr Ala Gln Ala Arg Phe Pro Ile Tyr Gly  
 145 150 155 160  
 Ile Pro Asp Asp Phe Ile Ser Val Pro Leu Pro Ala Gly Leu Arg Ser  
 165 170 175  
 30 Gly Lys Ala Leu Val Arg Ile Arg Gln Thr Gly Lys Asn Ser Gly Thr  
 180 185 190  
 Ile Asp Asn Thr Gly Gly Thr His Thr Ala Asp Leu Ser Arg Phe Pro  
 195 200 205  
 35 Ile Thr Ala Arg Thr Thr Ala Ile Lys Gly Arg Phe Glu Gly Ser Arg  
 210 215 220  
 Phe Leu Pro Tyr His Thr Arg Asn Gln Ile Asn Gly Gly Ala Leu Asp  
 225 230 235 240  
 40 Gly Lys Ala Pro Ile Leu Gly Tyr Ala Glu Asp Pro Val Glu Leu Phe  
 245 250 255  
 Phe Met His Ile Gln Gly Ser Gly Arg Leu Lys Thr Pro Ser Gly Lys  
 260 265 270  
 Tyr Ile Arg Ile Gly Tyr Ala Asp Lys Asn Glu His Pro Tyr Val Ser  
 275 280 285  
 45 Ile Gly Arg Tyr Met Ala Asp Lys Gly Tyr Leu Lys Leu Gly Gln Thr  
 290 295 300  
 Ser Met Gln Gly Ile Lys Ser Tyr Met Arg Gln Asn Pro Gln Arg Leu  
 305 310 315 320  
 50 Ala Glu Val Leu Gly Gln Asn Pro Ser Tyr Ile Phe Phe Arg Glu Leu  
 325 330 335  
 Ala Gly Ser Ser Asn Asp Gly Pro Val Gly Ala Leu Gly Thr Pro Leu  
 340 345 350  
 55 Met Gly Glu Tyr Ala Gly Ala Val Asp Arg His Tyr Ile Thr Leu Gly  
 355 360 365

## EP 1 790 660 A2

Ala Pro Leu Phe Val Ala Thr Ala His Pro Val Thr Arg Lys Ala Leu  
 370 375 380 385 390 395 400  
 Asn Arg Leu Ile Met Ala Gln Asp Thr Gly Ser Ala Ile Lys Gly Ala  
 5

Val Arg Val Asp Tyr Phe Trp Gly Tyr Gly Asp Glu Ala Gly Glu Leu  
 405 410 415  
 Ala Gly Lys Gln Lys Thr Thr Gly Tyr Val Trp Gln Leu Leu Pro Asn  
 10 420 425 430  
 Gly Met Lys Pro Glu Tyr Arg Pro  
 435 440  
 <210> 4  
 <211> 58  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 907-2.pep  
 20 <400> 4  
 Glu Arg Arg Arg Leu Leu Val Asn Ile Gln Tyr Glu Ser Ser Arg Ala  
 1 5 10 15  
 Gly Leu Asp Thr Gln Ile Val Leu Gly Leu Ile Glu Val Glu Ser Ala  
 25 30  
 Phe Arg Gln Tyr Ala Ile Ser Gly Val Gly Ala Arg Gly Leu Met Gln  
 35 40 45  
 Val Met Pro Phe Trp Lys Asn Tyr Ile Gly  
 30 50 55  
 <210> 5  
 <211> 60  
 <212> PRT  
 <213> Artificial Sequence  
 35 <220>  
 <223> Escherichia coli  
 <400> 5  
 Glu Arg Phe Pro Leu Ala Tyr Asn Asp Leu Phe Lys Arg Tyr Thr Ser  
 1 5 10 15  
 Gly Lys Glu Ile Pro Gln Ser Tyr Ala Met Ala Ile Ala Arg Gln Glu  
 40 20 25 30  
 Ser Ala Trp Asn Pro Lys Val Lys Ser Pro Val Gly Ala Ser Gly Leu  
 35 40 45  
 Met Gln Ile Met Pro Gly Thr Ala Thr His Thr Val  
 45 50 55 60  
 <210> 6  
 <211> 120  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 922.pep  
 <400> 6  
 Val Ala Gln Lys Tyr Gly Val Pro Ala Glu Leu Ile Val Ala Val Ile  
 55 1 5 10 15

Gly Ile Glu Thr Asn Tyr Gly Lys Asn Thr Gly Ser Phe Arg Val Ala  
 20 25 30

5 Asp Ala Leu Ala Thr Leu Gly Phe Asp Tyr Pro Arg Arg Ala Gly Phe  
 35 40 45

Phe Gln Lys Glu Leu Val Glu Leu Leu Lys Leu Ala Lys Glu Glu Gly  
 50 55 60

10 Gly Asp Val Phe Ala Phe Lys Gly Ser Tyr Ala Gly Ala Met Gly Met  
 65 70 75 80

Pro Gln Phe Met Pro Ser Ser Tyr Arg Lys Trp Ala Val Asp Tyr Asp  
 85 90 95

15 Gly Asp Gly His Arg Asp Ile Trp Gly Asn Val Gly Asp Val Ala Ala  
 100 105 110

Ser Val Ala Asn Tyr Met Lys Gln  
 115 120

20 <210> 7  
 <211> 119  
 <212> PRT  
 <213> Artificial Sequence

25 <220>  
 <223> Escherichia coli

30 <400> 7  
 Ala Trp Gln Val Tyr Gly Val Pro Pro Glu Ile Ile Val Gly Ile Ile  
 1 5 10 15

Gly Val Glu Thr Arg Trp Gly Arg Val Met Gly Lys Thr Arg Ile Leu  
 20 25 30

35 Asp Ala Leu Ala Thr Leu Ser Phe Asn Tyr Pro Arg Arg Ala Glu Tyr  
 35 40 45

Phe Ser Gly Glu Leu Glu Thr Phe Leu Leu Met Ala Arg Asp Glu Gln  
 50 55 60

40 Asp Asp Pro Leu Asn Leu Lys Gly Ser Phe Ala Gly Ala Met Gly Tyr  
 65 70 75 80

Gly Gln Phe Met Pro Ser Ser Tyr Lys Gln Tyr Ala Val Asp Phe Ser  
 85 90 95

45 Gly Asp Gly His Ile Asn Leu Trp Asp Pro Val Asp Ala Ile Gly Ser  
 100 105 110

Val Ala Asn Tyr Phe Lys Ala  
 115

50 <210> 8  
 <211> 194  
 <212> PRT  
 <213> Artificial Sequence

55 <220>  
 <223> 919.pep

<400> 8  
 Ala Leu Asp Gly Lys Ala Pro Ile Leu Gly Tyr Ala Glu Asp Pro Val  
 1 5 10 15

Glu Leu Phe Phe Met His Ile Gln Gly Ser Gly Arg Leu Lys Thr Pro  
 20 25 30

Ser Gly Lys Tyr Ile Arg Ile Gly Tyr Ala Asp Lys Asn Glu His Pro  
 35 40 45  
 5 Tyr Val Ser Ile Gly Arg Tyr Met Ala Asp Lys Gly Tyr Leu Lys Leu  
 50 55 60  
 Gly Gln Thr Ser Met Gln Gly Ile Lys Ser Tyr Met Arg Gln Asn Pro  
 65 70 75 80  
 10 Gln Arg Leu Ala Glu Val Leu Gly Gln Asn Pro Ser Tyr Ile Phe Phe  
 85 90 95  
 Arg Glu Leu Ala Gly Ser Ser Asn Asp Gly Pro Val Gly Ala Leu Gly  
 100 105 110  
 15 Thr Pro Leu Met Gly Glu Tyr Ala Gly Ala Val Asp Arg His Tyr Ile  
 115 120 125  
 Thr Leu Gly Ala Pro Leu Phe Val Ala Thr Ala His Pro Val Thr Arg  
 130 135 140  
 20 Lys Ala Leu Asn Arg Leu Ile Met Ala Gln Asp Thr Gly Ser Ala Ile  
 145 150 155 160  
 Lys Gly Ala Val Arg Val Asp Tyr Phe Trp Gly Tyr Gly Asp Glu Ala  
 165 170 175  
 25 Gly Glu Leu Ala Gly Lys Gln Lys Thr Thr Gly Tyr Val Trp Gln Leu  
 180 185 190  
 Leu Pro  
 30 <210> 9  
 <211> 196  
 <212> PRT  
 <213> Escherichia coli  
 <400> 9  
 Ala Leu Ser Asp Lys Tyr Ile Leu Ala Tyr Ser Asn Ser Leu Met Asp  
 1 5 10 15  
 35 Asn Phe Ile Met Asp Val Gln Gly Ser Gly Tyr Ile Asp Phe Gly Asp  
 20 25 30  
 Gly Ser Pro Leu Asn Phe Phe Ser Tyr Ala Gly Lys Asn Gly His Ala  
 35 40 45  
 40 Tyr Arg Ser Ile Gly Lys Val Leu Ile Asp Arg Gly Glu Val Lys Lys  
 50 55 60  
 Glu Asp Met Ser Met Gln Ala Ile Arg His Trp Gly Glu Thr His Ser  
 65 70 75 80  
 45 Glu Ala Glu Val Arg Glu Leu Leu Glu Gln Asn Pro Ser Phe Val Phe  
 85 90 95  
 Phe Lys Pro Gln Ser Phe Ala Pro Val Lys Gly Ala Ser Ala Val Pro  
 100 105 110  
 50 Leu Val Gly Arg Ala Ser Val Ala Ser Asp Arg Ser Ile Ile Pro Pro  
 115 120 125  
 Gly Thr Thr Leu Leu Ala Glu Val Pro Leu Leu Asp Asn Asn Gly Lys  
 130 135 140  
 55 Phe Asn Gly Gln Tyr Glu Leu Arg Leu Met Val Ala Leu Asp Val Gly  
 145 150 155 160

Gly Ala Ile Lys Gly Gln His Phe Asp Ile Tyr Gln Gly Ile Gly Pro  
 165 170 175  
 5 Glu Ala Gly His Arg Ala Gly Trp Tyr Asn His Tyr Gly Arg Val Trp  
 180 185 190  
 Val Leu Lys Thr  
 195  
 10 <210> 10  
 <211> 28  
 <212> DNA  
 <213> Artificial Sequence  
 15 <220>  
 <223> Oligonucleotide  
 <400> 10 cgaagacccc gtcggtcttt ttttatg 28  
 <210> 11  
 <211> 28  
 20 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> Oligonucleotide  
 25 <400> 11 gtgcataaaa aaaagaccga cgggtct 28  
 <210> 12  
 <211> 25  
 <212> DNA  
 30 <213> Artificial Sequence  
 <220>  
 <223> Oligonucleotide  
 <400> 12 aacgcctcgc cggtgttttg ggtca 25  
 35 <210> 13  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
 40 <220>  
 <223> Oligonucleotide  
 <400> 13 tttgacccaa aacaccggcg aggcg 25  
 45 <210> 14  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence  
 50 <220>  
 <223> Oligonucleotide  
 <400> 14 tgccggcgca gtcggtcggc actaca 26  
 55 <210> 15  
 <211> 26  
 <212> DNA  
 <213> Artificial Sequence

	<220>			
	<223>	Oligonucleotide		
5	<400>	15		
	taatgttagtg	ccgaccgact	gcgccc	
	<210>	16		
	<211>	25		
	<212>	DNA		
10	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
15	<400>	16		
	tgatttaggtt	ggtagcgcg	ttcccg	
	<210>	17		
	<211>	25		
	<212>	DNA		
	<213>	Artificial Sequence		
20	<220>			
	<223>	Oligonucleotide		
	<400>	17		
	ggcggaaacgc	gctacccacc	tcaat	
25	<210>	18		
	<211>	34		
	<212>	DNA		
	<213>	Artificial Sequence		
30	<220>			
	<223>	Oligonucleotide		
	<400>	18		
	ccggaattct	tatgaaaaaaaa	atcatcttcg	ccgc
35	<210>	19		
	<211>	32		
	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
40	<400>	19		
	gccccaagctt	ttattgtttg	gctgcctcga	tt
	<210>	20		
	<211>	37		
	<212>	DNA		
45	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
50	<400>	20		
	ccggaattct	tatgtcgccc	gatgttaaat	cggcgga
	<210>	21		
	<211>	32		
	<212>	DNA		
	<213>	Artificial Sequence		
55	<220>			
	<223>	Oligonucleotide		

	<400> 21		
	gccccaaagctt tcaatcctgc tctttttgc cg		32
5	<210> 22		
	<211> 34		
	<212> DNA		
	<213> Artificial Sequence		
10	<220>		
	<223> oligonucleotide		
	<400> 22		
	ccggaaattct tatgagccaa gatatggcgg cagt		34
15	<210> 23		
	<211> 32		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
	<223> oligonucleotide		
20	<400> 23		
	gccccaaagctt tcaatcctgc tctttttgc cg		32
	<210> 24		
	<211> 34		
	<212> DNA		
25	<213> Artificial Sequence		
	<220>		
	<223> oligonucleotide		
	<400> 24		
30	ccggaaattct tatgtccgcc gaatccgcaa atca		34
	<210> 25		
	<211> 32		
	<212> DNA		
	<213> Artificial Sequence		
35	<220>		
	<223> oligonucleotide		
	<400> 25		
	gccccaaagctt tcaatcctgc tctttttgc cg		32
40	<210> 26		
	<211> 36		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
45	<223> oligonucleotide		
	<400> 26		
	ccggaaattct tatgggaagg gttgatttgg ctaatg		36
	<210> 27		
	<211> 32		
	<212> DNA		
50	<213> Artificial Sequence		
	<220>		
	<223> oligonucleotide		
55	<400> 27		
	gccccaaagctt tcaatcctgc tctttttgc cg		32

	<210>	28	
	<211>	36	
5	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
10	<400>	28	
	ccggaattct tatgtcagat ttggcaaacg attctt		36
	<210>	29	
	<211>	35	
	<212>	DNA	
15	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
20	<400>	29	
	gcccaagctt ttacgtatca tatttcacgt gcttc		35
	<210>	30	
	<211>	37	
	<212>	DNA	
25	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
30	<400>	30	
	ccggaattct tatgtcgccc gatgttaaat cggcgga		37
	<210>	31	
	<211>	35	
	<212>	DNA	
35	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
40	<400>	31	
	gcccaagctt ttacgtatca tatttcacgt gcttc		35
	<210>	32	
	<211>	36	
	<212>	DNA	
45	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
50	<400>	32	
	ccggaattct tatgcaaagc aagagcatcc aaacct		36
	<210>	33	
	<211>	30	
	<212>	DNA	
55	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	33	
	gcccaagctt ttacgggcgg tattcgggct		30
	<210>	34	
	<211>	29	
	<212>	DNA	

	<213> Artificial Sequence	
5	<220>	
	<223> oligonucleotide	
	<400> 34	
	ccggaattca tatgaaacac tttccatcc	29
10	<210> 35	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
15	<400> 35	
	gcccaagctt ttaccactcg taattgac	28
	<210> 36	
	<211> 29	
	<212> DNA	
20	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
	<400> 36	
25	ccggaattca tatggccaca agcgacgac	29
	<210> 37	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> oligonucleotide	
	<400> 37	
	gcccaagctt ttaccactcg taattgac	28
35	<210> 38	
	<211> 29	
	<212> DNA	
	<213> Artificial Sequence	
40	<220>	
	<223> oligonucleotide	
	<400> 38	
	ccggaattct tatgaaacac tttccatcc	29
45	<210> 39	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
50	<400> 39	
	gcccaagctt tcaacccacg ttgttaaggtt g	31
	<210> 40	
	<211> 30	
	<212> DNA	
55	<213> Artificial Sequence	
	<220>	

	<223> Oligonucleotide	
5	<400> 40 ccggaattct tatggccaca aacgacgacg	30
	<210> 41	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> Oligonucleotide	
	<400> 41 gcccaagctt tcaacccacg ttgttaaggtt g	31
15	<210> 42	
	<211> 34	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223> Oligonucleotide	
	<400> 42 ccggaattct tatggccacc tacaaagtgg acga	34
25	<210> 43	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
30	<400> 43 gcccaagctt ttattgtttg gctgcctcga tt	32
	<210> 44	
	<211> 32	
	<212> DNA	
35	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 44 cgcgatccg ctagccccga tgttaaatcg gc	32
40	<210> 45	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
45	<220>	
	<223> Oligonucleotide	
	<400> 45 cccgctcgag tcaatcctgc tctttttgc c	31
50	<210> 46	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
55	<220>	
	<223> Oligonucleotide	
	<400> 46	

	cgcgatccg ctagccaaga tatggcggca gt	32
5	<210> 47 <211> 32 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
10	<400> 47 cgcgatccg ctagcgccga atccgcaa at ca	32
	<210> 48 <211> 32 <212> DNA <213> Artificial Sequence	
15	<220> <223> oligonucleotide	
	<400> 48 cgcgctagcg gaagggttga tttggctaat gg	32
20	<210> 49 <211> 34 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
25	<400> 49 gggaattcca tatgggcatt tcccgcaaaa tatc	34
30	<210> 50 <211> 32 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
35	<400> 50 cccgctcgag ttacgtatca tatttcacgt gc	32
	<210> 51 <211> 34 <212> DNA <213> Artificial Sequence	
40	<220> <223> oligonucleotide	
	<400> 51 gggaattcca tatgggcatt tcccgcaaaa tatc	34
45	<210> 52 <211> 33 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
50	<400> 52 cccgctcgag ttattctatg ccttgtgcgg cat	33
	<210> 53	

	<211>	32	
	<212>	DNA	
	<213>	Artificial Sequence	
5	<220>		
	<223>	Oligonucleotide	
	<400>	53	
		cgccggatccc atatggccac aagcgacgac ga	32
10	<210>	54	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
15	<223>	Oligonucleotide	
	<400>	54	
		cccgctcgag ttaccactcg taattgac	28
20	<210>	55	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
25	<223>	Oligonucleotide	
	<400>	55	
		cgccggatccc atatggccac aaacgacg	28
	<210>	56	
	<211>	35	
	<212>	DNA	
30	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	56	
35		cccgctcgag tcatttagca atattatctt tgttc	35
	<210>	57	
	<211>	33	
	<212>	DNA	
	<213>	Artificial Sequence	
40	<220>		
	<223>	Oligonucleotide	
	<400>	57	
		cgccggatccc atatgaaagc aaacagtgcc gac	33
45	<210>	58	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
50	<223>	Oligonucleotide	
	<400>	58	
		cccgctcgag ttaccactcg taattgac	28
	<210>	59	
	<211>	28	
	<212>	DNA	
55	<213>	Artificial Sequence	

	<220>			
	<223>	oligonucleotide		
5	<400>	59		
	cgcgatccc	atatggccac	aaacgacg	
	<210>	60		
	<211>	29		
	<212>	DNA		
10	<213>	Artificial Sequence		
	<220>			
	<223>	oligonucleotide		
	<400>	60		
15	cccgctcgag	ttaaccacg	ttgttaagg	
	<210>	61		
	<211>	33		
	<212>	DNA		
	<213>	Artificial Sequence		
20	<220>			
	<223>	oligonucleotide		
	<400>	61		
	cgcgatccc	atatgatgaa	acactttcca	tcc
25	<210>	62		
	<211>	29		
	<212>	DNA		
	<213>	Artificial Sequence		
30	<220>			
	<223>	oligonucleotide		
	<400>	62		
	cccgctcgag	ttaaccacg	ttgttaagg	
35	<210>	63		
	<211>	28		
	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
	<223>	oligonucleotide		
40	<400>	63		
	cgcgatccc	atatggccac	aaacgacg	
	<210>	64		
	<211>	32		
	<212>	DNA		
45	<213>	Artificial Sequence		
	<220>			
	<223>	oligonucleotide		
	<400>	64		
50	cccgctcgag	tcagtctgac	actgttttat	cc
	<210>	65		
	<211>	32		
	<212>	DNA		
	<213>	Artificial Sequence		
55	<220>			
	<223>	oligonucleotide		

	<400> 65	
	cgcggatccg ctagccccga tgttaaatcg gc	32
5	<210> 66	
	<211> 27	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> oligonucleotide	
	<400> 66	
	cccgctcgag ttacgggcgg tattcgg	27
15	<210> 67	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
20	<400> 67	
	cgcggatccg ctagccccga tgttaaatcg gc	32
	<210> 68	
	<211> 32	
	<212> DNA	
25	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
	<400> 68	
30	cccgctcgag ttacgtatca tatttcacgt gc	32
	<210> 69	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
35	<220>	
	<223> oligonucleotide	
	<400> 69	
	cgcggatccg ctagccccga tgttaaatcg gc	32
40	<210> 70	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
45	<220>	
	<223> Oligonucleotide	
	<400> 70	
	cccgctcgag ttaccactcg taattgac	28
50	<210> 71	
	<211> 1457	
	<212> PRT	
	<213> Neisseria meningitidis	
	<400> 71	
55	Met Lys Thr Thr Asp Lys Arg Thr Thr Glu Thr His Arg Lys Ala Pro	
	1 5 10 15	
	Lys Thr Gly Arg Ile Arg Phe Ser Pro Ala Tyr Leu Ala Ile Cys Leu	

	20	25	30
5	Ser Phe Gly Ile Leu Pro Gln Ala Trp Ala Gly His Thr Tyr Phe Gly 35 40 45		
	Ile Asn Tyr Gln Tyr Tyr Arg Asp Phe Ala Glu Asn Lys Gly Lys Phe 50 55 60		
10	Ala Val Gly Ala Lys Asp Ile Glu Val Tyr Asn Lys Lys Gly Glu Leu 65 70 75 80		
	Val Gly Lys Ser Met Thr Lys Ala Pro Met Ile Asp Phe Ser Val Val 85 90 95		
15	Ser Arg Asn Gly Val Ala Ala Leu Val Gly Asp Gln Tyr Ile Val Ser 100 105 110		
	Val Ala His Asn Gly Gly Tyr Asn Asn Val Asp Phe Gly Ala Glu Gly 115 120 125		
20	Arg Asn Pro Asp Gln His Arg Phe Thr Tyr Lys Ile Val Lys Arg Asn 130 135 140		
	Asn Tyr Lys Ala Gly Thr Lys Gly His Pro Tyr Gly Gly Asp Tyr His 145 150 155 160		
	Met Pro Arg Leu His Lys Phe Val Thr Asp Ala Glu Pro Val Glu Met 165 170 175		
25	Thr Ser Tyr Met Asp Gly Arg Lys Tyr Ile Asp Gln Asn Asn Tyr Pro 180 185 190		
	Asp Arg Val Arg Ile Gly Ala Gly Arg Gln Tyr Trp Arg Ser Asp Glu 195 200 205		
30	Asp Glu Pro Asn Asn Arg Glu Ser Ser Tyr His Ile Ala Ser Ala Tyr 210 215 220		
	Ser Trp Leu Val Gly Gly Asn Thr Phe Ala Gln Asn Gly Ser Gly Gly 225 230 235 240		
35	Gly Thr Val Asn Leu Gly Ser Glu Lys Ile Lys His Ser Pro Tyr Gly 245 250 255		
	Phe Leu Pro Thr Gly Gly Ser Phe Gly Asp Ser Gly Ser Pro Met Phe 260 265 270		
40	Ile Tyr Asp Ala Gln Lys Gln Lys Trp Leu Ile Asn Gly Val Leu Gln 275 280 285		
	Thr Gly Asn Pro Tyr Ile Gly Lys Ser Asn Gly Phe Gln Leu Val Arg 290 295 300		
45	Lys Asp Trp Phe Tyr Asp Glu Ile Phe Ala Gly Asp Thr His Ser Val 305 310 315 320		
	Phe Tyr Glu Pro Arg Gln Asn Gly Lys Tyr Ser Phe Asn Asp Asp Asn 325 330 335		
50	Asn Gly Thr Gly Lys Ile Asn Ala Lys His Glu His Asn Ser Leu Pro 340 345 350		
	Asn Arg Leu Lys Thr Arg Thr Val Gln Leu Phe Asn Val Ser Leu Ser 355 360 365		
55	Glu Thr Ala Arg Glu Pro Val Tyr His Ala Ala Gly Gly Val Asn Ser 370 375 380		
	Tyr Arg Pro Arg Leu Asn Asn Gly Glu Asn Ile Ser Phe Ile Asp Glu		

## EP 1 790 660 A2

	385	390	395	400
5	Gly Lys Gly Glu Leu Ile Leu Thr Ser Asn Ile Asn Gln Gly Ala Gly 405 410 415			
	Gly Leu Tyr Phe Gln Gly Asp Phe Thr Val Ser Pro Glu Asn Asn Glu 420 425 430			
10	Thr Trp Gln Gly Ala Gly Val His Ile Ser Glu Asp Ser Thr Val Thr 435 440 445			
	Trp Lys Val Asn Gly Val Ala Asn Asp Arg Leu Ser Lys Ile Gly Lys 450 455 460			
15	Gly Thr Leu His Val Gln Ala Lys Gly Glu Asn Gln Gly Ser Ile Ser 465 470 475 480			
	Val Gly Asp Gly Thr Val Ile Leu Asp Gln Gln Ala Asp Asp Lys Gly 485 490 495			
20	Lys Lys Gln Ala Phe Ser Glu Ile Gly Leu Val Ser Gly Arg Gly Thr 500 505 510			
	Val Gln Leu Asn Ala Asp Asn Gln Phe Asn Pro Asp Lys Leu Tyr Phe 515 520 525			
25	Gly Phe Arg Gly Gly Arg Leu Asp Leu Asn Gly His Ser Leu Ser Phe 530 535 540			
	His Arg Ile Gln Asn Thr Asp Glu Gly Ala Met Ile Val Asn His Asn 545 550 555 560			
	Gln Asp Lys Glu Ser Thr Val Thr Ile Thr Gly Asn Lys Asp Ile Ala 565 570 575			
30	Thr Thr Gly Asn Asn Asn Ser Leu Asp Ser Lys Lys Glu Ile Ala Tyr 580 585 590			
	Asn Gly Trp Phe Gly Glu Lys Asp Thr Thr Lys Thr Asn Gly Arg Leu 595 600 605			
35	Asn Leu Val Tyr Gln Pro Ala Ala Glu Asp Arg Thr Leu Leu Leu Ser 610 615 620			
	Gly Gly Thr Asn Leu Asn Gln Asn Ile Thr Gln Thr Asn Gly Lys Leu 625 630 635 640			
40	Phe Phe Ser Gly Arg Pro Thr Pro His Ala Tyr Asn His Leu Asn Asp 645 650 655			
	His Trp Ser Gln Lys Glu Gly Ile Pro Arg Gly Glu Ile Val Trp Asp 660 665 670			
45	Asn Asp Trp Ile Asn Arg Thr Phe Lys Ala Glu Asn Phe Gln Ile Lys 675 680 685			
	Gly Gly Gln Ala Val Val Ser Arg Asn Val Ala Lys Val Lys Gly Asp 690 695 700			
50	Trp His Leu Ser Asn His Ala Gln Ala Val Phe Gly Val Ala Pro His 705 710 715 720			
	Gln Ser His Thr Ile Cys Thr Arg Ser Asp Trp Thr Gly Leu Thr Asn 725 730 735			
55	Cys Val Glu Lys Thr Ile Thr Asp Asp Lys Val Ile Ala Ser Leu Thr 740 745 750			
	Lys Thr Asp Ile Ser Gly Asn Val Asp Leu Ala Asp His Ala His Leu			

## EP 1 790 660 A2

	755	760	765
5	Asn Leu Thr Gly Leu Ala Thr Leu Asn Gly Asn Leu Ser Ala Asn Gly 770 775 780		
	Asp Thr Arg Tyr Thr Val Ser His Asn Ala Thr Gln Asn Gly Asn Leu 785 790 795 800		
10	Ser Leu Val Gly Asn Ala Gln Ala Thr Phe Asn Gln Ala Thr Leu Asn 805 810 815		
	Gly Asn Thr Ser Ala Ser Gly Asn Ala Ser Phe Asn Leu Ser Asp His 820 825 830		
15	Ala Val Gln Asn Gly Ser Leu Thr Leu Ser Gly Asn Ala Lys Ala Asn 835 840 845		
	Val Ser His Ser Ala Leu Asn Gly Asn Val Ser Leu Ala Asp Lys Ala 850 855 860		
20	Val Phe His Phe Glu Ser Ser Arg Phe Thr Gly Gln Ile Ser Gly Gly 865 870 875 880		
	Lys Asp Thr Ala Leu His Leu Lys Asp Ser Glu Trp Thr Leu Pro Ser 885 890 895		
25	Gly Thr Glu Leu Gly Asn Leu Asn Leu Asp Asn Ala Thr Ile Thr Leu 900 905 910		
	Asn Ser Ala Tyr Arg His Asp Ala Ala Gly Ala Gln Thr Gly Ser Ala 915 920 925		
30	Thr Asp Ala Pro Arg Arg Arg Ser Arg Arg Ser Arg Arg Ser Leu Leu 930 935 940		
	Ser Val Thr Pro Pro Thr Ser Val Glu Ser Arg Phe Asn Thr Leu Thr 945 950 955 960		
35	Val Asn Gly Lys Leu Asn Gly Gln Gly Thr Phe Arg Phe Met Ser Glu 965 970 975		
	Leu Phe Gly Tyr Arg Ser Asp Lys Leu Lys Leu Ala Glu Ser Ser Glu 980 985 990		
40	Gly Thr Tyr Thr Leu Ala Val Asn Asn Thr Gly Asn Glu Pro Ala Ser 995 1000 1005		
	Leu Glu Gln Leu Thr Val Val Glu Gly Lys Asp Asn Lys Pro Leu Ser 1010 1015 1020		
45	Glu Asn Leu Asn Phe Thr Leu Gln Asn Glu His Val Asp Ala Gly Ala 1025 1030 1035 1040		
	Trp Arg Tyr Gln Leu Ile Arg Lys Asp Gly Glu Phe Arg Leu His Asn 1045 1050 1055		
50	Pro Val Lys Glu Gln Glu Leu Ser Asp Lys Leu Gly Lys Ala Glu Ala 1060 1065 1070		
	Lys Lys Gln Ala Glu Lys Asp Asn Ala Gln Ser Leu Asp Ala Leu Ile 1075 1080 1085		
55	Ala Ala Gly Arg Asp Ala Val Glu Lys Thr Glu Ser Val Ala Glu Pro 1090 1095 1100		
	Ala Arg Gln Ala Gly Gly Glu Asn Val Gly Ile Met Gln Ala Glu Glu 1105 1110 1115 1120		
	Glu Lys Lys Arg Val Gln Ala Asp Lys Asp Thr Ala Leu Ala Lys Gln		

	1125	1130	1135
5	Arg Glu Ala Glu Thr Arg Pro Ala Thr Thr Ala Phe Pro Arg Ala Arg 1140 1145 1150		
	Arg Ala Arg Arg Asp Leu Pro Gln Leu Gln Pro Gln Pro Gln 1155 1160 1165		
10	Pro Gln Arg Asp Leu Ile Ser Arg Tyr Ala Asn Ser Gly Leu Ser Glu 1170 1175 1180		
	Phe Ser Ala Thr Leu Asn Ser Val Phe Ala Val Gln Asp Glu Leu Asp 1185 1190 1195 1200		
15	Arg Val Phe Ala Glu Asp Arg Arg Asn Ala Val Trp Thr Ser Gly Ile 1205 1210 1215		
	Arg Asp Thr Lys His Tyr Arg Ser Gln Asp Phe Arg Ala Tyr Arg Gln 1220 1225 1230		
20	Gln Thr Asp Leu Arg Gln Ile Gly Met Gln Lys Asn Leu Gly Ser Gly 1235 1240 1245		
	Arg Val Gly Ile Leu Phe Ser His Asn Arg Thr Glu Asn Thr Phe Asp 1250 1255 1260		
25	Asp Gly Ile Gly Asn Ser Ala Arg Leu Ala His Gly Ala Val Phe Gly 1265 1270 1275 1280		
	Gln Tyr Gly Ile Asp Arg Phe Tyr Ile Gly Ile Ser Ala Gly Ala Gly 1285 1290 1295		
	Phe Ser Ser Gly Ser Leu Ser Asp Gly Ile Gly Gly Lys Ile Arg Arg 1300 1305 1310		
30	Arg Val Leu His Tyr Gly Ile Gln Ala Arg Tyr Arg Ala Gly Phe Gly 1315 1320 1325		
	Gly Phe Gly Ile Glu Pro His Ile Gly Ala Thr Arg Tyr Phe Val Gln 1330 1335 1340		
35	Lys Ala Asp Tyr Arg Tyr Glu Asn Val Asn Ile Ala Thr Pro Gly Leu 1345 1350 1355 1360		
	Ala Phe Asn Arg Tyr Arg Ala Gly Ile Lys Ala Asp Tyr Ser Phe Lys 1365 1370 1375		
40	Pro Ala Gln His Ile Ser Ile Thr Pro Tyr Leu Ser Leu Ser Tyr Thr 1380 1385 1390		
	Asp Ala Ala Ser Gly Lys Val Arg Thr Arg Val Asn Thr Ala Val Leu 1395 1400 1405		
45	Ala Gln Asp Phe Gly Lys Thr Arg Ser Ala Glu Trp Gly Val Asn Ala 1410 1415 1420		
	Glu Ile Lys Gly Phe Thr Leu Ser Leu His Ala Ala Ala Ala Lys Gly 1425 1430 1435 1440		
50	Pro Gln Leu Glu Ala Gln His Ser Ala Gly Ile Lys Leu Gly Tyr Arg 1445 1450 1455		
	Trp		
55	<210> 72 <211> 21 <212> PRT <213> Escherichia coli		

<400> 72  
 Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala  
 1 5 10 15  
 5 Thr Val Ala Gln Ala  
 20  
 <210> 73  
 <211> 1439  
 <212> PRT  
 10 <213> Neisseria meningitidis  
 <400> 73  
 Met Lys Lys Thr Ala Ile Ala Ile Ala Val Ala Leu Ala Gly Phe Ala  
 1 5 10 15  
 15 Thr Val Ala Gln Ala Ala Ser Ala Gly His Thr Tyr Phe Gly Ile Asn  
 20 25 30  
 Tyr Gln Tyr Tyr Arg Asp Phe Ala Glu Asn Lys Gly Lys Phe Ala Val  
 35 40 45  
 20 Gly Ala Lys Asp Ile Glu Val Tyr Asn Lys Lys Gly Glu Leu Val Gly  
 50 55 60  
 Lys Ser Met Thr Lys Ala Pro Met Ile Asp Phe Ser Val Val Ser Arg  
 65 70 75 80  
 25 Asn Gly Val Ala Ala Leu Val Gly Asp Gln Tyr Ile Val Ser Val Ala  
 85 90 95  
 His Asn Gly Gly Tyr Asn Asn Val Asp Phe Gly Ala Glu Gly Arg Asn  
 100 105 110  
 30 Pro Asp Gln His Arg Phe Thr Tyr Lys Ile Val Lys Arg Asn Asn Tyr  
 115 120 125  
 Lys Ala Gly Thr Lys Gly His Pro Tyr Gly Gly Asp Tyr His Met Pro  
 130 135 140  
 35 Arg Leu His Lys Phe Val Thr Asp Ala Glu Pro Val Glu Met Thr Ser  
 145 150 155 160  
 Tyr Met Asp Gly Arg Lys Tyr Ile Asp Gln Asn Asn Tyr Pro Asp Arg  
 165 170 175  
 40 Val Arg Ile Gly Ala Gly Arg Gln Tyr Trp Arg Ser Asp Glu Asp Glu  
 180 185 190  
 Pro Asn Asn Arg Glu Ser Ser Tyr His Ile Ala Ser Ala Tyr Ser Trp  
 195 200 205  
 45 Leu Val Gly Gly Asn Thr Phe Ala Gln Asn Gly Ser Gly Gly Thr  
 210 215 220  
 Val Asn Leu Gly Ser Glu Lys Ile Lys His Ser Pro Tyr Gly Phe Leu  
 225 230 235 240  
 50 Pro Thr Gly Gly Ser Phe Gly Asp Ser Gly Ser Pro Met Phe Ile Tyr  
 245 250 255  
 Asp Ala Gln Lys Gln Lys Trp Leu Ile Asn Gly Val Leu Gln Thr Gly  
 260 265 270  
 55 Asn Pro Tyr Ile Gly Lys Ser Asn Gly Phe Gln Leu Val Arg Lys Asp  
 275 280 285  
 Trp Phe Tyr Asp Glu Ile Phe Ala Gly Asp Thr His Ser Val Phe Tyr

## EP 1 790 660 A2

	290	295	300
5	Glu Pro Arg Gln Asn Gly Lys Tyr Ser Phe Asn Asp Asp Asn Asn Gly 305 310 315 320		
	Thr Gly Lys Ile Asn Ala Lys His Glu His Asn Ser Leu Pro Asn Arg 325 330 335		
10	Leu Lys Thr Arg Thr Val Gln Leu Phe Asn Val Ser Leu Ser Glu Thr 340 345 350		
	Ala Arg Glu Pro Val Tyr His Ala Ala Gly Gly Val Asn Ser Tyr Arg 355 360 365		
15	Pro Arg Leu Asn Asn Gly Glu Asn Ile Ser Phe Ile Asp Glu Gly Lys 370 375 380		
	Gly Glu Leu Ile Leu Thr Ser Asn Ile Asn Gln Gly Ala Gly Gly Leu 385 390 395 400		
	Tyr Phe Gln Gly Asp Phe Thr Val Ser Pro Glu Asn Asn Glu Thr Trp 405 410 415		
20	Gln Gly Ala Gly Val His Ile Ser Glu Asp Ser Thr Val Thr Trp Lys 420 425 430		
	Val Asn Gly Val Ala Asn Asp Arg Leu Ser Lys Ile Gly Lys Gly Thr 435 440 445		
25	Leu His Val Gln Ala Lys Gly Glu Asn Gln Gly Ser Ile Ser Val Gly 450 455 460		
	Asp Gly Thr Val Ile Leu Asp Gln Gln Ala Asp Asp Lys Gly Lys Lys 465 470 475 480		
30	Gln Ala Phe Ser Glu Ile Gly Leu Val Ser Gly Arg Gly Thr Val Gln 485 490 495		
	Leu Asn Ala Asp Asn Gln Phe Asn Pro Asp Lys Leu Tyr Phe Gly Phe 500 505 510		
35	Arg Gly Gly Arg Leu Asp Leu Asn Gly His Ser Leu Ser Phe His Arg 515 520 525		
	Ile Gln Asn Thr Asp Glu Gly Ala Met Ile Val Asn His Asn Gln Asp 530 535 540		
40	Lys Glu Ser Thr Val Thr Ile Thr Gly Asn Lys Asp Ile Ala Thr Thr 545 550 555 560		
	Gly Asn Asn Asn Ser Leu Asp Ser Lys Lys Glu Ile Ala Tyr Asn Gly 565 570 575		
45	Trp Phe Gly Glu Lys Asp Thr Thr Lys Thr Asn Gly Arg Leu Asn Leu 580 585 590		
	Val Tyr Gln Pro Ala Ala Glu Asp Arg Thr Leu Leu Leu Ser Gly Gly 595 600 605		
50	Thr Asn Leu Asn Gly Asn Ile Thr Gln Thr Asn Gly Lys Leu Phe Phe 610 615 620		
	Ser Gly Arg Pro Thr Pro His Ala Tyr Asn His Leu Asn Asp His Trp 625 630 635 640		
55	Ser Gln Lys Glu Gly Ile Pro Arg Gly Glu Ile Val Trp Asp Asn Asp 645 650 655		
	Trp Ile Asn Arg Thr Phe Lys Ala Glu Asn Phe Gln Ile Lys Gly Gly		

## EP 1 790 660 A2

	660	665	670
5	Gln Ala Val Val Ser Arg Asn Val Ala Lys Val Lys Gly Asp Trp His 675 680 685		
	Leu Ser Asn His Ala Gln Ala Val Phe Gly Val Ala Pro His Gln Ser 690 695 700		
10	His Thr Ile Cys Thr Arg Ser Asp Trp Thr Gly Leu Thr Asn Cys Val 705 710 715 720		
	Glu Lys Thr Ile Thr Asp Asp Lys Val Ile Ala Ser Leu Thr Lys Thr 725 730 735		
15	Asp Ile Ser Gly Asn Val Asp Leu Ala Asp His Ala His Leu Asn Leu 740 745 750		
	Thr Gly Leu Ala Thr Leu Asn Gly Asn Leu Ser Ala Asn Gly Asp Thr 755 760 765		
20	Arg Tyr Thr Val Ser His Asn Ala Thr Gln Asn Gly Asn Leu Ser Leu 770 775 780		
	Val Gly Asn Ala Gln Ala Thr Phe Asn Gln Ala Thr Leu Asn Gly Asn 785 790 795 800		
	Thr Ser Ala Ser Gly Asn Ala Ser Phe Asn Leu Ser Asp His Ala Val 805 810 815		
25	Gln Asn Gly Ser Leu Thr Leu Ser Gly Asn Ala Lys Ala Asn Val Ser 820 825 830		
	His Ser Ala Leu Asn Gly Asn Val Ser Leu Ala Asp Lys Ala Val Phe 835 840 845		
30	His Phe Glu Ser Ser Arg Phe Thr Gly Gln Ile Ser Gly Gly Lys Asp 850 855 860		
	Thr Ala Leu His Leu Lys Asp Ser Glu Trp Thr Leu Pro Ser Gly Thr 865 870 875 880		
35	Glu Leu Gly Asn Leu Asn Leu Asp Asn Ala Thr Ile Thr Leu Asn Ser 885 890 895		
	Ala Tyr Arg His Asp Ala Ala Gly Ala Gln Thr Gly Ser Ala Thr Asp 900 905 910		
40	Ala Pro Arg Arg Arg Ser Arg Arg Ser Arg Arg Ser Leu Leu Ser Val 915 920 925		
	Thr Pro Pro Thr Ser Val Glu Ser Arg Phe Asn Thr Leu Thr Val Asn 930 935 940		
45	Gly Lys Leu Asn Gly Gln Gly Thr Phe Arg Phe Met Ser Glu Leu Phe 945 950 955 960		
	Gly Tyr Arg Ser Asp Lys Leu Lys Leu Ala Glu Ser Ser Glu Gly Thr 965 970 975		
50	Tyr Thr Leu Ala Val Asn Asn Thr Gly Asn Glu Pro Ala Ser Leu Glu 980 985 990		
	Gln Leu Thr Val Val Glu Gly Lys Asp Asn Lys Pro Leu Ser Glu Asn 995 1000 1005		
55	Leu Asn Phe Thr Leu Gln Asn Glu His Val Asp Ala Gly Ala Trp Arg 1010 1015 1020		
	Tyr Gln Leu Ile Arg Lys Asp Gly Glu Phe Arg Leu His Asn Pro Val		

## EP 1 790 660 A2

	1025	1030	1035	1040
5	Lys Glu Gln Glu Leu Ser Asp Lys Leu Gly Lys Ala Glu Ala Lys Lys 1045		1050	1055
	Gln Ala Glu Lys Asp Asn Ala Gln Ser Leu Asp Ala Leu Ile Ala Ala 1060	1065		1070
10	Gly Arg Asp Ala Val Glu Lys Thr Glu Ser Val Ala Glu Pro Ala Arg 1075	1080		1085
	Gln Ala Gly Gly Glu Asn Val Gly Ile Met Gln Ala Glu Glu Glu Lys 1090	1095		1100
15	Lys Arg Val Gln Ala Asp Lys Asp Thr Ala Leu Ala Lys Gln Arg Glu 1105	1110	1115	1120
	Ala Glu Thr Arg Pro Ala Thr Thr Ala Phe Pro Arg Ala Arg Arg Ala 1125	1130		1135
20	Arg Arg Asp Leu Pro Gln Leu Gln Pro Gln Pro Gln Pro Gln Pro Gln 1140	1145		1150
	Arg Asp Leu Ile Ser Arg Tyr Ala Asn Ser Gly Leu Ser Glu Phe Ser 1155	1160		1165
25	Ala Thr Leu Asn Ser Val Phe Ala Val Gln Asp Glu Leu Asp Arg Val 1170	1175		1180
	Phe Ala Glu Asp Arg Arg Asn Ala Val Trp Thr Ser Gly Ile Arg Asp 1185	1190	1195	1200
	Thr Lys His Tyr Arg Ser Gln Asp Phe Arg Ala Tyr Arg Gln Gln Thr 1205	1210		1215
30	Asp Leu Arg Gln Ile Gly Met Gln Lys Asn Leu Gly Ser Gly Arg Val 1220	1225		1230
	Gly Ile Leu Phe Ser His Asn Arg Thr Glu Asn Thr Phe Asp Asp Gly 1235	1240		1245
35	Ile Gly Asn Ser Ala Arg Leu Ala His Gly Ala val Phe Gly Gln Tyr 1250	1255		1260
	Gly Ile Asp Arg Phe Tyr Ile Gly Ile Ser Ala Gly Ala Gly Phe Ser 1265	1270	1275	1280
40	Ser Gly Ser Leu Ser Asp Gly Ile Gly Gly Lys Ile Arg Arg Arg Val 1285	1290		1295
	Leu His Tyr Gly Ile Gln Ala Arg Tyr Arg Ala Gly Phe Gly Gly Phe 1300	1305		1310
45	Gly Ile Glu Pro His Ile Gly Ala Thr Arg Tyr Phe Val Gln Lys Ala 1315	1320		1325
	Asp Tyr Arg Tyr Glu Asn Val Asn Ile Ala Thr Pro Gly Leu Ala Phe 1330	1335		1340
50	Asn Arg Tyr Arg Ala Gly Ile Lys Ala Asp Tyr Ser Phe Lys Pro Ala 1345	1350	1355	1360
	Gln His Ile Ser Ile Thr Pro Tyr Leu Ser Leu Ser Tyr Thr Asp Ala 1365	1370		1375
55	Ala Ser Gly Lys Val Arg Thr Arg Val Asn Thr Ala Val Leu Ala Gln 1380	1385		1390
	Asp Phe Gly Lys Thr Arg Ser Ala Glu Trp Gly Val Asn Ala Glu Ile			

	1395	1400	1405
5	Lys Gly Phe Thr Leu Ser Leu His Ala Ala Ala Ala Lys Gly Pro Gln 1410 1415 1420		
	Leu Glu Ala Gln His Ser Ala Gly Ile Lys Leu Gly Tyr Arg Trp 1425 1430 1435		
10	<210> 74 <211> 164 <212> PRT <213> Neisseria meningitidis		
	<400> 74 Met Lys Lys Asn Ile Leu Glu Phe Trp Val Gly Leu Phe Val Leu Ile 1 5 10 15		
15	Gly Ala Ala Ala Val Ala Phe Leu Ala Phe Arg Val Ala Gly Gly Ala 20 25 30		
	Ala Phe Gly Gly Ser Asp Lys Thr Tyr Ala Val Tyr Ala Asp Phe Gly 35 40 45		
20	Asp Ile Gly Gly Leu Lys Val Asn Ala Pro Val Lys Ser Ala Gly Val 50 55 60		
	Leu Val Gly Arg Val Gly Ala Ile Gly Leu Asp Pro Lys Ser Tyr Gln 65 70 75 80		
25	Ala Arg Val Arg Leu Asp Leu Asp Gly Lys Tyr Gln Phe Ser Ser Asp 85 90 95		
	Val Ser Ala Gln Ile Leu Thr Ser Gly Leu Leu Gly Glu Gln Tyr Ile 100 105 110		
30	Gly Leu Gln Gln Gly Gly Asp Thr Glu Asn Leu Ala Ala Gly Asp Thr 115 120 125		
	Ile Ser Val Thr Ser Ser Ala Met Val Leu Glu Asn Leu Ile Gly Lys 130 135 140		
35	Phe Met Thr Ser Phe Ala Glu Lys Asn Ala Asp Gly Gly Asn Ala Glu 145 150 155 160		
	Lys Ala Ala Glu		
40	<210> 75 <211> 21 <212> PRT <213> Erwinia carotovora		
	<400> 75 Met Lys Tyr Leu Leu Pro Thr Ala Ala Ala Gly Leu Leu Leu Ala Ala 1 5 10 15		
	Gln Pro Ala Met Ala 20		
45	<210> 76 <211> 608 <212> PRT <213> Neisseria meningitidis ORF46		
	<400> 76 Leu Gly Ile Ser Arg Lys Ile Ser Leu Ile Leu Ser Ile Leu Ala Val 1 5 10 15		
50	Cys Leu Pro Met His Ala His Ala Ser Asp Leu Ala Asn Asp Ser Phe		
55			

	20	25	30	
5	Ile Arg Gln Val Leu Asp Arg Gln His Phe Glu Pro Asp Gly Lys Tyr 35	40	45	
	His Leu Phe Gly Ser Arg Gly Glu Leu Ala Glu Arg Ser Gly His Ile 50	55	60	
10	Gly Leu Gly Lys Ile Gln Ser His Gln Leu Gly Asn Leu Met Ile Gln 65	70	75	80
	Gln Ala Ala Ile Lys Gly Asn Ile Gly Tyr Ile Val Arg Phe Ser Asp 85	90	95	
15	His Gly His Glu Val His Ser Pro Phe Asp Asn His Ala Ser His Ser 100	105	110	
	Asp Ser Asp Glu Ala Gly Ser Pro Val Asp Gly Phe Ser Leu Tyr Arg 115	120	125	
20	Ile His Trp Asp Gly Tyr Glu His His Pro Ala Asp Gly Tyr Asp Gly 130	135	140	
	Pro Gln Gly Gly Tyr Pro Ala Pro Lys Gly Ala Arg Asp Ile Tyr 145	150	155	160
25	Ser Tyr Asp Ile Lys Gly Val Ala Gln Asn Ile Arg Leu Asn Leu Thr 165	170	175	
	Asp Asn Arg Ser Thr Gly Gln Arg Leu Ala Asp Arg Phe His Asn Ala 180	185	190	
30	Gly Ser Met Leu Thr Gln Gly Val Gly Asp Gly Phe Lys Arg Ala Thr 195	200	205	
	Arg Tyr Ser Pro Glu Leu Asp Arg Ser Gly Asn Ala Ala Glu Ala Phe 210	215	220	
35	Asn Gly Thr Ala Asp Ile Val Lys Asn Ile Ile Gly Ala Ala Gly Glu 225	230	235	240
	Ile Val Gly Ala Gly Asp Ala Val Gln Gly Ile Ser Glu Gly Ser Asn 245	250	255	
40	Ile Ala Val Met His Gly Leu Gly Leu Leu Ser Thr Glu Asn Lys Met 260	265	270	
	Ala Arg Ile Asn Asp Leu Ala Asp Met Ala Gln Leu Lys Asp Tyr Ala 275	280	285	
45	Ala Ala Ala Ile Arg Asp Trp Ala Val Gln Asn Pro Asn Ala Ala Gln 290	295	300	
	Gly Ile Glu Ala Val Ser Asn Ile Phe Met Ala Ala Ile Pro Ile Lys 305	310	315	320
50	Gly Ile Gly Ala Val Arg Gly Lys Tyr Gly Leu Gly Gly Ile Thr Ala 325	330	335	
	His Pro Ile Lys Arg Ser Gln Met Gly Ala Ile Ala Leu Pro Lys Gly 340	345	350	
55	Lys Ser Ala Val Ser Asp Asn Phe Ala Asp Ala Ala Tyr Ala Lys Tyr 355	360	365	
	Pro Ser Pro Tyr His Ser Arg Asn Ile Arg Ser Asn Leu Glu Gln Arg 370	375	380	
	Tyr Gly Lys Glu Asn Ile Thr Ser Ser Thr Val Pro Pro Ser Asn Gly			

## EP 1 790 660 A2

	385	390	395	400
5	Lys Asn Val Lys Leu Ala Asp Gln Arg His Pro Lys Thr Gly Val Pro 405 410 415			
	Phe Asp Gly Lys Gly Phe Pro Asn Phe Glu Lys His Val Lys Tyr Asp 420 425 430			
10	Thr Lys Leu Asp Ile Gln Glu Leu Ser Gly Gly Ile Pro Lys Ala 435 440 445			
	Lys Pro Val Ser Asp Ala Lys Pro Arg Trp Glu Val Asp Arg Lys Leu 450 455 460			
15	Asn Lys Leu Thr Thr Arg Glu Gln Val Glu Lys Asn Val Gln Glu Ile 465 470 475 480			
	Arg Asn Gly Asn Lys Asn Ser Asn Phe Ser Gln His Ala Gln Leu Glu 485 490 495			
20	Arg Glu Ile Asn Lys Leu Lys Ser Ala Asp Glu Ile Asn Phe Ala Asp 500 505 510			
	Gly Met Gly Lys Phe Thr Asp Ser Met Asn Asp Lys Ala Phe Ser Arg 515 520 525			
25	Leu Val Lys Ser Val Lys Glu Asn Gly Phe Thr Asn Pro Val Val Glu 530 535 540			
	Tyr Val Glu Ile Asn Gly Lys Ala Tyr Ile Val Arg Gly Asn Asn Arg 545 550 555 560			
	Val Phe Ala Ala Glu Tyr Leu Gly Arg Ile His Glu Leu Lys Phe Lys 565 570 575			
30	Lys Val Asp Phe Pro Val Pro Asn Thr Ser Trp Lys Asn Pro Thr Asp 580 585 590			
	Val Leu Asn Glu Ser Gly Asn Val Lys Arg Pro Arg Tyr Arg Ser Lys 595 600 605			
35	<210> 77 <211> 584 <212> PRT <213> Artificial Sequence			
40	<220> <223> ORF46-2			
	<400> 77 Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg Gln 1 5 10 15			
45	His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg Gly Glu 20 25 30			
	Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln Ser His 35 40 45			
50	Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly Asn Ile 50 55 60			
	Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu Val His Ser Pro 65 70 75 80			
55	Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly Ser Pro 85 90 95			

Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu His  
 100 105 110  
 5 His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Tyr Pro Ala  
 115 120 125  
 Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val Ala  
 130 135 140  
 10 Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln Arg  
 145 150 155 160  
 Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly Val  
 165 170 175  
 15 Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp Arg  
 180 185 190  
 Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val Lys  
 195 200 205  
 20 Asn Ile Ile Gly Ala Ala Glu Ile Val Gly Ala Gly Asp Ala Val  
 210 215 220  
 Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu Gly  
 225 230 235 240  
 25 Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala Asp  
 245 250 255  
 Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ile Arg Asp Trp Ala  
 260 265 270  
 30 Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn Ile  
 275 280 285  
 Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly Lys  
 290 295 300  
 35 Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln Met  
 305 310 315 320  
 Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn Phe  
 325 330 335  
 Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg Asn  
 340 345 350  
 40 Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr Ser  
 355 360 365  
 Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp Gln  
 370 375 380  
 45 Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro Asn  
 385 390 395 400  
 Phe Glu Lys His Val Lys Tyr Asp Thr Lys Leu Asp Ile Gln Glu Leu  
 405 410 415  
 50 Ser Gly Gly Ile Pro Lys Ala Lys Pro Val Ser Asp Ala Lys Pro  
 420 425 430  
 Arg Trp Glu Val Asp Arg Lys Leu Asn Lys Leu Thr Thr Arg Glu Gln  
 435 440 445  
 55 Val Glu Lys Asn Val Gln Glu Ile Arg Asn Gly Asn Lys Asn Ser Asn  
 450 455 460

Phe Ser Gln His Ala Gln Leu Glu Arg Glu Ile Asn Lys Leu Lys Ser  
 465 470 475 480  
 5 Ala Asp Glu Ile Asn Phe Ala Asp Gly Met Gly Lys Phe Thr Asp Ser  
 485 490 495  
 Met Asn Asp Lys Ala Phe Ser Arg Leu Val Lys Ser Val Lys Glu Asn  
 500 505 510  
 10 Gly Phe Thr Asn Pro Val Val Glu Tyr Val Glu Ile Asn Gly Lys Ala  
 515 520 525  
 Tyr Ile Val Arg Gly Asn Asn Arg Val Phe Ala Ala Glu Tyr Leu Gly  
 530 535 540  
 15 Arg Ile His Glu Leu Lys Phe Lys Lys Val Asp Phe Pro Val Pro Asn  
 545 550 555 560  
 Thr Ser Trp Lys Asn Pro Thr Asp Val Leu Asn Glu Ser Gly Asn Val  
 565 570 575  
 20 Lys Arg Pro Arg Tyr Arg Ser Lys  
 580  
 <210> 78  
 <211> 364  
 <212> PRT  
 <213> Neisseria meningitidis  
 25 <400> 78  
 Met Ser Met Lys His Phe Pro Ala Lys Val Leu Thr Thr Ala Ile Leu  
 1 5 10 15  
 Ala Thr Phe Cys Ser Gly Ala Leu Ala Ala Thr Ser Asp Asp Asp Val  
 20 25 30  
 30 Lys Lys Ala Ala Thr Val Ala Ile Val Ala Ala Tyr Asn Asn Gly Gln  
 35 40 45  
 Glu Ile Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile Gly Glu  
 50 55 60  
 35 Asp Gly Thr Ile Thr Gln Lys Asp Ala Thr Ala Ala Asp Val Glu Ala  
 65 70 75 80  
 Asp Asp Phe Lys Gly Leu Lys Lys Val Val Thr Asn Leu Thr  
 85 90 95  
 40 Lys Thr Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val Lys Ala  
 100 105 110  
 Ala Glu Ser Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp Thr Asp  
 115 120 125  
 45 Ala Ala Leu Ala Asp Thr Asp Ala Ala Leu Asp Glu Thr Thr Asn Ala  
 130 135 140  
 Leu Asn Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu Thr Lys  
 145 150 155 160  
 50 Thr Asn Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala Asp Thr  
 165 170 175  
 Val Asp Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp Ser Leu Asp  
 180 185 190  
 55 Glu Thr Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala Asn Glu Ala  
 195 200 205

Lys Gln Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala Lys Val Lys  
 210 215 220  
 5 Ala Ala Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Ala Gly Thr Ala  
 225 230 235 240  
 Asn Thr Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys Val Thr Asp  
 245 250 255  
 10 Ile Lys Ala Asp Ile Ala Thr Asn Lys Ala Asp Ile Ala Lys Asn Ser  
 260 265 270  
 Ala Arg Ile Asp Ser Leu Asp Lys Asn Val Ala Asn Leu Arg Lys Glu  
 275 280 285  
 15 Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu Phe Gln  
 290 295 300  
 Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val Gly Gly Tyr  
 305 310 315 320  
 20 Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg Phe Thr Glu  
 325 330 335  
 Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser Ser Gly Ser  
 340 345 350  
 25 Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp  
 355 360  
 <210> 79  
 <211> 427  
 <212> PRT  
 <213> Neisseria meningitidis  
 30 <400> 79  
 Met Phe Glu Arg Ser Val Ile Ala Met Ala Cys Ile Phe Ala Leu Ser  
 1 5 10 15  
 Ala Cys Gly Gly Gly Gly Ser Pro Asp Val Lys Ser Ala Asp  
 20 25 30  
 35 Thr Leu Ser Lys Pro Ala Ala Pro Val Val Ala Glu Lys Glu Thr Glu  
 35 40 45  
 Val Lys Glu Asp Ala Pro Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro  
 50 55 60  
 40 Ser Thr Gln Gly Ser Gln Asp Met Ala Ala Val Ser Ala Glu Asn Thr  
 65 70 75 80  
 Gly Asn Gly Gly Ala Ala Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu  
 85 90 95  
 45 Gly Pro Gln Asn Asp Met Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln  
 100 105 110  
 Thr Gly Asn Asn Gln Pro Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser  
 115 120 125  
 50 Asn Pro Ala Pro Ala Asn Gly Gly Ser Asn Phe Gly Arg Val Asp Leu  
 130 135 140  
 Ala Asn Gly Val Leu Ile Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr  
 145 150 155 160  
 55 His Cys Lys Gly Asp Ser Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu  
 165 170 175

Ala Pro Ser Lys Ser Glu Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile  
 180 185 190  
 5 Glu Lys Tyr Lys Lys Asp Gly Lys Ser Asp Lys Phe Thr Asn Leu Val  
 195 200 205  
 Ala Thr Ala Val Gln Ala Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr  
 210 215 220  
 10 Lys Asp Lys Ser Ala Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala  
 225 230 235 240  
 Arg Ser Arg Arg Ser Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn  
 245 250 255  
 15 Gln Ala Asp Thr Leu Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly  
 260 265 270  
 His Ser Gly Asn Ile Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr  
 275 280 285  
 20 Tyr Gly Ala Glu Lys Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln  
 290 295 300  
 Gly Glu Pro Ala Lys Gly Glu Met Leu Ala Gly Thr Ala Val Tyr Asn  
 305 310 315 320  
 25 Gly Glu Val Leu His Phe His Thr Glu Asn Gly Arg Pro Tyr Pro Thr  
 325 330 335  
 Arg Gly Arg Phe Ala Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp  
 340 345 350  
 30 Gly Ile Ile Asp Ser Gly Asp Asp Leu His Met Gly Thr Gln Lys Phe  
 355 360 365  
 Lys Ala Ala Ile Asp Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn  
 370 375 380  
 35 Gly Gly Asp Val Ser Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu  
 385 390 395 400  
 Val Ala Gly Lys Tyr Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly  
 405 410 415  
 40 Phe Gly Val Phe Ala Gly Lys Lys Glu Gln Asp  
 420 425  
 <210> 80  
 <211> 410  
 <212> PRT  
 <213> Artificial Sequence  
 45 <220>  
 <223> 287untagged  
 <400> 80  
 Cys Gly Gly Gly Gly Gly Ser Pro Asp Val Lys Ser Ala Asp Thr  
 1 5 10 15  
 50 Leu Ser Lys Pro Ala Ala Pro Val Val Ala Glu Lys Glu Thr Glu Val  
 20 25 30  
 Lys Glu Asp Ala Pro Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser  
 35 40 45  
 55 Thr Gln Gly Ser Gln Asp Met Ala Ala Val Ser Ala Glu Asn Thr Gly  
 50 55 60

Asn Gly Gly Ala Ala Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly  
 65 70 75 80

5 Pro Gln Asn Asp Met Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln Thr  
 85 90 95

Gly Asn Asn Gln Pro Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser Asn  
 100 105 110

10 Pro Ala Pro Ala Asn Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala  
 115 120 125

Asn Gly Val Leu Ile Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His  
 130 135 140

15 Cys Lys Gly Asp Ser Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu Ala  
 145 150 155 160

Pro Ser Lys Ser Glu Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile Glu  
 165 170 175

20 Lys Tyr Lys Lys Asp Gly Lys Ser Asp Lys Phe Thr Asn Leu Val Ala  
 180 185 190

Thr Ala Val Gln Ala Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr Lys  
 195 200

25 Asp Lys Ser Ala Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala Arg  
 210 215 220

Ser Arg Arg Ser Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln  
 225 230 235 240

Ala Asp Thr Leu Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly His  
 245 250 255

30 Ser Gly Asn Ile Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr  
 260 265 270

Gly Ala Glu Lys Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly  
 275 280 285

35 Glu Pro Ala Lys Gly Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly  
 290 295 300

Glu Val Leu His Phe His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg  
 305 310 315 320

40 Gly Arg Phe Ala Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly  
 325 330 335

Ile Ile Asp Ser Gly Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys  
 340 345 350

45 Ala Ala Ile Asp Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly  
 355 360 365

Gly Gly Asp Val Ser Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu Val  
 370 375 380

50 Ala Gly Lys Tyr Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe  
 385 390 395 400

Gly Val Phe Ala Gly Lys Lys Glu Gln Asp  
 405 410

55 <210> 81  
 <211> 9  
 <212> PRT

5           <213> Artificial Sequence  
 <220>  
 <223> 920L N-terminal  
 <400> 81  
 His Arg Val Trp Val Glu Thr Ala His  
 1                 5  
 10           <210> 82  
 <211> 16  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 953L N-terminal  
 15           <400> 82  
 Ala Thr Tyr Lys Val Asp Glu Tyr His Ala Asn Ala Arg Phe Ala Phe  
 1                 5                     10                 15  
 20           <210> 83  
 <211> 9  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 519.1L N-terminal  
 25           <400> 83  
 Met Glu Phe Phe Ile Ile Leu Leu Ala  
 1                 5  
 30           <210> 84  
 <211> 488  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> deltaG287  
 35           <400> 84  
 Met Phe Lys Arg Ser Val Ile Ala Met Ala Cys Ile Phe Ala Leu Ser  
 1                 5                     10                 15  
 Ala Cys Gly Gly Gly Gly Ser Pro Asp Val Lys Ser Ala Asp  
 40           20                 25                 30  
 Thr Leu Ser Lys Pro Ala Ala Pro Val Val Ser Glu Lys Glu Thr Glu  
 35           35                 40                 45  
 Ala Lys Glu Asp Ala Pro Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro  
 45           50                 55                 60  
 Ser Ala Gln Gly Ser Gln Asp Met Ala Ala Val Ser Glu Glu Asn Thr  
 65           70                 75                 80  
 Gly Asn Gly Gly Ala Val Thr Ala Asp Asn Pro Lys Asn Glu Asp Glu  
 50           85                 90                 95  
 Val Ala Gln Asn Asp Met Pro Gln Asn Ala Ala Gly Thr Asp Ser Ser  
 100           105                 110  
 Thr Pro Asn His Thr Pro Asp Pro Asn Met Leu Ala Gly Asn Met Glu  
 115           120                 125  
 Asn Gln Ala Thr Asp Ala Gly Glu Ser Ser Gln Pro Ala Asn Gln Pro  
 55           130                 135                 140

## EP 1 790 660 A2

Asp Met Ala Asn Ala Ala Asp Gly Met Gln Gly Asp Asp Pro Ser Ala  
 145 150 155 160  
 5 Gly Gly Gln Asn Ala Gly Asn Thr Ala Ala Gln Gly Ala Asn Gln Ala  
 165 170 175  
 Gly Asn Asn Gln Ala Ala Gly Ser Ser Asp Pro Ile Pro Ala Ser Asn  
 180 185 190  
 10 Pro Ala Pro Ala Asn Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala  
 195 200 205  
 Asn Gly Val Leu Ile Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His  
 210 215 220  
 15 Cys Lys Gly Asp Ser Cys Ser Gly Asn Asn Phe Leu Asp Glu Glu Val  
 225 230 235 240  
 Gln Leu Lys Ser Glu Phe Glu Lys Leu Ser Asp Ala Asp Lys Ile Ser  
 245 250 255  
 20 Asn Tyr Lys Lys Asp Gly Lys Asn Asp Lys Phe Val Gly Leu Val Ala  
 260 265 270  
 Asp Ser Val Gln Met Lys Gly Ile Asn Gln Tyr Ile Ile Phe Tyr Lys  
 275 280 285  
 25 Pro Lys Pro Thr Ser Phe Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg  
 290 295 300  
 Arg Ser Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp  
 305 310 315 320  
 30 Thr Leu Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly  
 325 330 335  
 Asn Ile Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala  
 340 345 350  
 35 Glu Lys Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro  
 355 360 365  
 Ala Lys Gly Glu Met Leu Ala Gly Ala Ala Val Tyr Asn Gly Glu Val  
 370 375 380  
 40 Leu His Phe His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg Gly Arg  
 385 390 395 400  
 Phe Ala Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile  
 405 410 415  
 45 Asp Ser Gly Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala  
 420 425 430  
 Ile Asp Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Ser Gly  
 435 440 445  
 Asp Val Ser Gly Lys Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly  
 450 455 460  
 50 Lys Tyr Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val  
 465 470 475 480  
 Phe Ala Gly Lys Lys Glu Gln Asp  
 485  
 55 <210> 85  
 <211> 712

<212> PRT  
 <213> Artificial Sequence

5           <220>  
             <223> TBP2

<400> 85  
 Met Asn Asn Pro Leu Val Asn Gln Ala Ala Met Val Leu Pro Val Phe  
 1                 5                 10                 15

10           Leu Leu Ser Ala Cys Leu Gly Gly Gly Ser Phe Asp Leu Asp Ser  
                20                 25                 30

15           Val Asp Thr Glu Ala Pro Arg Pro Ala Pro Lys Tyr Gln Asp Val Phe  
                35                 40                 45

20           Ser Glu Lys Pro Gln Ala Gln Lys Asp Gln Gly Gly Tyr Gly Phe Ala  
                50                 55                 60

25           Met Arg Leu Lys Arg Arg Asn Trp Tyr Pro Gln Ala Lys Glu Asp Glu  
                65                 70                 75                 80

30           Val Lys Leu Asp Glu Ser Asp Trp Glu Ala Thr Gly Leu Pro Asp Glu  
                85                 90                 95

35           Pro Lys Glu Leu Pro Lys Arg Gln Lys Ser Val Ile Glu Lys Val Glu  
                100                 105                 110

40           Thr Asp Ser Asp Asn Asn Ile Tyr Ser Ser Pro Tyr Leu Lys Pro Ser  
                115                 120                 125

45           Asn His Gln Asn Gly Asn Thr Gly Asn Gly Ile Asn Gln Pro Lys Asn  
                130                 135                 140

50           Gln Ala Lys Asp Tyr Glu Asn Phe Lys Tyr Val Tyr Ser Gly Trp Phe  
                145                 150                 155                 160

55           Tyr Lys His Ala Lys Arg Glu Phe Asn Leu Lys Val Glu Pro Lys Ser  
                165                 170                 175

60           Ala Lys Asn Gly Asp Asp Gly Tyr Ile Phe Tyr His Gly Lys Glu Pro  
                180                 185                 190

65           Ser Arg Gln Leu Pro Ala Ser Gly Lys Ile Thr Tyr Lys Gly Val Trp  
                195                 200                 205

70           His Phe Ala Thr Asp Thr Lys Lys Gly Gln Lys Phe Arg Glu Ile Ile  
                210                 215                 220

75           Gln Pro Ser Lys Ser Gln Gly Asp Arg Tyr Ser Gly Phe Ser Gly Asp  
                225                 230                 235                 240

80           Asp Gly Glu Glu Tyr Ser Asn Lys Asn Lys Ser Thr Leu Thr Asp Gly  
                245                 250                 255

85           Gln Glu Gly Tyr Gly Phe Thr Ser Asn Leu Glu Val Asp Phe His Asn  
                260                 265                 270

90           Lys Lys Leu Thr Gly Lys Leu Ile Arg Asn Asn Ala Asn Thr Asp Asn  
                275                 280                 285

95           Asn Gln Ala Thr Thr Thr Gln Tyr Tyr Ser Leu Glu Ala Gln Val Thr  
                290                 295                 300

100           Gly Asn Arg Phe Asn Gly Lys Ala Thr Ala Thr Asp Lys Pro Gln Gln  
                305                 310                 315                 320

105           Asn Ser Glu Thr Lys Glu His Pro Phe Val Ser Asp Ser Ser Leu  
                325                 330                 335

Ser Gly Gly Phe Phe Gly Pro Gln Gly Glu Glu Leu Gly Phe Arg Phe  
 340 345 350  
 5 Leu Ser Asp Asp Gln Lys Val Ala Val Val Gly Ser Ala Lys Thr Lys  
 355 360 365  
 Asp Lys Pro Ala Asn Gly Asn Thr Ala Ala Ala Ser Gly Gly Thr Asp  
 370 375 380 385  
 10 Ala Ala Ala Ser Asn Gly Ala Ala Gly Thr Ser Ser Glu Asn Gly Lys  
 390 395 400  
 Leu Thr Thr Val Leu Asp Ala Val Glu Leu Lys Leu Gly Asp Lys Glu  
 405 410 415  
 15 Val Gln Lys Leu Asp Asn Phe Ser Asn Ala Ala Gln Leu Val Val Asp  
 420 425 430  
 Gly Ile Met Ile Pro Leu Leu Pro Glu Ala Ser Glu Ser Gly Asn Asn  
 435 440 445  
 20 Gln Ala Asn Gln Gly Thr Asn Gly Gly Thr Ala Phe Thr Arg Lys Phe  
 450 455 460  
 Asp His Thr Pro Glu Ser Asp Lys Lys Asp Ala Gln Ala Gly Thr Gln  
 465 470 475 480  
 25 Thr Asn Gly Ala Gln Thr Ala Ser Asn Thr Ala Gly Asp Thr Asn Gly  
 485 490 495  
 Lys Thr Lys Thr Tyr Glu Val Glu Val Cys Cys Ser Asn Leu Asn Tyr  
 500 505 510  
 30 Leu Lys Tyr Gly Met Leu Thr Arg Lys Asn Ser Lys Ser Ala Met Gln  
 515 520 525  
 Ala Gly Glu Ser Ser Ser Gln Ala Asp Ala Lys Thr Glu Gln Val Glu  
 530 535 540  
 35 Gln Ser Met Phe Leu Gln Gly Glu Arg Thr Asp Glu Lys Glu Ile Pro  
 545 550 555 560  
 Ser Glu Gln Asn Ile Val Tyr Arg Gly Ser Trp Tyr Gly Tyr Ile Ala  
 565 570 575  
 40 Asn Asp Lys Ser Thr Ser Trp Ser Gly Asn Ala Ser Asn Ala Thr Ser  
 580 585 590  
 Gly Asn Arg Ala Glu Phe Thr Val Asn Phe Ala Asp Lys Lys Ile Thr  
 595 600 605  
 45 Gly Thr Leu Thr Ala Asp Asn Arg Gln Glu Ala Thr Phe Thr Ile Asp  
 610 615 620  
 Gly Asn Ile Lys Asp Asn Gln Phe Glu Gly Thr Ala Lys Thr Ala Glu  
 625 630 635 640  
 50 Ser Gly Phe Asp Leu Asp Gln Ser Asn Thr Thr Arg Thr Pro Lys Ala  
 645 650 655  
 Tyr Ile Thr Asp Ala Lys Val Gln Gly Gly Phe Tyr Gly Pro Lys Ala  
 660 665 670  
 Glu Glu Leu Gly Gly Trp Phe Ala Tyr Pro Gly Asp Lys Gln Thr Lys  
 675 680 685  
 55 Asn Ala Thr Asn Ala Ser Gly Asn Ser Ser Ala Thr Val Val Phe Gly  
 690 695 700

Ala Lys Arg Gln Gln Pro Val Arg  
 705  
 5        <210> 86  
 <211> 274  
 <212> PRT  
 <213> Artificial Sequence  
 10      <220>  
 <223> 741  
 15      <400> 86  
 Val Asn Arg Thr Ala Phe Cys Cys Leu Ser Leu Thr Thr Ala Leu Ile  
 1                    5                    10                    15  
 Leu Thr Ala Cys Ser Ser Gly Gly Gly Val Ala Ala Asp Ile Gly  
 20                    25                    30  
 Ala Gly Leu Ala Asp Ala Leu Thr Ala Pro Leu Asp His Lys Asp Lys  
 35                    40                    45  
 20      Gly Leu Gln Ser Leu Thr Leu Asp Gln Ser Val Arg Lys Asn Glu Lys  
 50                    55                    60  
 Leu Lys Leu Ala Ala Gln Gly Ala Glu Lys Thr Tyr Gly Asn Gly Asp  
 65                    70                    75                    80  
 25      Ser Leu Asn Thr Gly Lys Leu Lys Asn Asp Lys Val Ser Arg Phe Asp  
 85                    90                    95  
 Phe Ile Arg Gln Ile Glu Val Asp Gly Gln Leu Ile Thr Leu Glu Ser  
 100                  105                  110  
 30      Gly Glu Phe Gln Val Tyr Lys Gln Ser His Ser Ala Leu Thr Ala Phe  
 115                  120                  125  
 Gln Thr Glu Gln Ile Gln Asp Ser Glu His Ser Gly Lys Met Val Ala  
 130                  135                  140  
 35      Lys Arg Gln Phe Arg Ile Gly Asp Ile Ala Gly Glu His Thr Ser Phe  
 145                  150                  155                  160  
 Asp Lys Leu Pro Glu Gly Arg Ala Thr Tyr Arg Gly Thr Ala Phe  
 165                  170                  175  
 40      Gly Ser Asp Asp Ala Gly Gly Lys Leu Thr Tyr Thr Ile Asp Phe Ala  
 180                  185                  190  
 Ala Lys Gln Gly Asn Gly Lys Ile Glu His Leu Lys Ser Pro Glu Leu  
 195                  200                  205  
 45      Asn Val Asp Leu Ala Ala Asp Ile Lys Pro Asp Gly Lys Arg His  
 210                  215                  220  
 Ala Val Ile Ser Gly Ser Val Leu Tyr Asn Gln Ala Glu Lys Gly Ser  
 225                  230                  235                  240  
 Tyr Ser Leu Gly Ile Phe Gly Gly Lys Ala Gln Glu Val Ala Gly Ser  
 245                  250                  255  
 50      Ala Glu Val Lys Thr Val Asn Gly Ile Arg His Ile Gly Leu Ala Ala  
 260                  265                  270  
 Lys Gln  
 55      <210> 87  
 <211> 1082

<212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 983  
 <400> 87  
 Met Arg Thr Thr Pro Thr Phe Pro Thr Lys Thr Phe Lys Pro Thr Ala  
 1 5 10 15  
 Met Ala Leu Ala Val Ala Thr Thr Leu Ser Ala Cys Leu Gly Gly  
 20 25 30  
 Gly Gly Gly Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile  
 35 40 45  
 Gly Ser Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr  
 50 55 60  
 Ala Gly Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala  
 65 70 75 80  
 Gly Arg Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala  
 85 90 95  
 Pro Pro Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala  
 100 105 110  
 Tyr Lys Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr  
 115 120 125  
 Gly Arg Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly  
 130 135 140  
 Ser Ile Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn  
 145 150 155 160  
 Glu Asn Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu  
 165 170 175  
 Asp Gly Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val  
 180 185 190  
 Ile Glu Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile  
 195 200 205  
 Gly His Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp  
 210 215 220  
 Gly Arg Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met  
 225 230 235 240  
 Asn Thr Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg  
 245 250 255  
 Asn Ala Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn  
 260 265 270  
 Ser Phe Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile  
 275 280 285  
 Ala Asn Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly  
 290 295 300  
 Gly Asp Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr  
 305 310 315 320  
 Gly Asn Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe  
 325 330 335

Ser Thr Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu  
 340 345 350  
 5 Pro Phe Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly  
 355 360 365  
 Val Asp Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro  
 370 375 380  
 10 Gly Thr Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala  
 385 390 395 400  
 Met Trp Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg  
 405 410 415  
 15 Thr Asn Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val  
 420 425 430  
 Thr Gly Thr Ala Ala Leu Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn  
 435 440 445  
 20 Asp Asn Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala  
 450 455 460  
 Val Gly Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys  
 465 470 475 480  
 25 Ala Met Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp  
 485 490 495  
 Thr Lys Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser  
 500 505 510  
 30 Gly Thr Gly Gly Leu Ile Lys Lys Gly Ser Gln Leu Gln Leu His  
 515 520 525  
 Gly Asn Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Ser Leu  
 530 535 540  
 35 Val Leu Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly  
 545 550 555 560  
 Ala Leu Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp  
 565 570 575  
 Gly Ile Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr  
 580 585 590  
 40 Val His Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr  
 595 600 605  
 Thr Arg Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly  
 610 615 620  
 45 Gly Lys Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn  
 625 630 635 640  
 Ser Thr Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln  
 645 650 655  
 50 Asp Tyr Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala  
 660 665 670  
 Ser Leu Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu  
 675 680 685  
 55 Ser Tyr Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala  
 690 695 700

Ala His Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln Gly Gly  
 705 710 715 720  
 5 Ser Asn Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu Ser Ser  
 725 730 735  
 Ala Thr Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr Asp Met  
 740 745 750  
 10 Pro Gly Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Val  
 755 760 765  
 Gln His Ala Asn Ala Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala  
 770 775 780  
 15 Ala Thr Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly  
 785 790 795 800  
 Arg Arg Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly  
 805 810 815  
 20 Leu Arg Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln  
 820 825 830  
 Gly Gly Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile  
 835 840 845  
 25 Ala Ala Lys Thr Gly Glu Asn Thr Thr Ala Ala Ala Thr Leu Gly Met  
 850 855 860  
 Gly Arg Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser  
 865 870 875 880  
 30 Ile Ser Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr  
 885 890 895  
 Leu Lys Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg  
 900 905 910  
 35 Ser Thr Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu  
 915 920 925  
 Met Gln Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr  
 930 935 940  
 40 Gly Asp Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys Gln  
 945 950 955 960  
 Asp Ala Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser  
 965 970 975  
 45 Leu Thr Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln  
 980 985 990  
 Pro Leu Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val Glu Arg  
 995 1000 1005  
 50 Asp Leu Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly Ala  
 1010 1015 1020  
 Thr Ala Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His Thr Arg  
 1025 1030 1035 1040  
 Leu Val Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly Trp Asn  
 1045 1050 1055  
 55 Gly Leu Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn His  
 1060 1065 1070

Ser Gly Arg Val Gly Val Gly Tyr Arg Phe  
1075 1080

<210> 88  
<211> 2505  
212

<212> DNA  
<213> Artificial Sequence

<220>  
<223> deltaG287-919

<400> 88

atggcttagcc	ccgatgttaa	atcgccggac	acgctgtcaa	aaccggccgc	tcctgttgg	60
gctaaaaaaag	agacagaggt	aaaagaagat	gcgccacagg	caggttctca	aggacagggc	120
gcgcacatcca	acaaggcag	ccaagatatg	cgccgagttt	cggcagaaaa	tacaggcaat	180
ggcggtgcgg	caacaacgga	caaacccaa	aatgaagacg	agggaccgca	aatatgatatg	240
ccgaaattt	ccgcgaat	cgcaatcaa	acagggaaac	accaaccgc	cgatttctca	300
gatcccggcc	ccgcgtcaaa	ccctgcacct	gcaatggcg	gtaccaat	tggaaagggtt	360
gatttggcta	atggcggtt	gattgatggg	ccgtcgcaaa	ataaacgtt	gaccactgt	420
aaaggcgatt	cttgtaatgg	tgataattta	ttgatgaag	aagcaccgtc	aaaatcagaa	480
tttggaaaatt	taaatgagtc	tgaacgaatt	gagaatata	agaaagatgg	gaaaagcgat	540
aaatttacta	atttgggtgc	gacagcagtt	caagctaattg	gaactaaca	atatgtcatc	600
atttataaag	acaagtccgc	ttcatcttca	tctgcgcgat	tcaggcgttc	tgcacggctg	660
aggaggctgc	ttccgtccga	gatggcgta	atccccgtca	atcaggcgga	tacgtctatt	720
gtcgatgggg	aaggcgtag	cctgacgggg	cattccggca	atatcttcgc	gccgaagggg	780
aattaccgg	atctgactta	cgggggcgaa	aaatttgcgg	cgccgatctgta	tgcctccctgt	840
gtgcaaggcg	aaccggcaaa	aggcgaat	cttgcggca	cgggcggtgta	caacggcgaa	900
gtgctgcatt	ttcatacgga	aaacggccgt	ccgtacccga	ctagaggcag	gttgcgcga	960
aaagtcgatt	tcggcagcaa	atctgtggac	ggcattatcg	acagcggcga	tgatttgcatt	1020
atgggtacgc	aaaaattcaa	agccgcctac	gatggaaacg	gtttaaggg	gacttggacg	1080
aaaaatggcg	gccccggatgt	ttccggaaagg	tttacggcc	cgcccgcgca	ggaagtggcg	1140
ggaaaatatac	gctatcgccc	gacagatgcg	gaaaaggcg	gattcgctgt	gttgcggcc	1200
aaaaaaagagc	aggatggatc	cgaggaggaa	ggatgccaa	gcaagagcat	ccaaacccctt	1260
ccgcaacccg	acacatccgt	catcaacggc	ccggaccggc	cggtcgcat	ccccgaccccc	1320
gccggaacg	cggtcgccgg	cgccggggcc	gtctataccg	ttgtacccgca	cctgtccctg	1380
ccccacttgg	cggcgcagga	tttcgcacaa	agccctgcaat	ccttccgcct	cggtcgcc	1440
aatttggaaa	accgccaagg	ctggcaggat	gtgtgcgc	aagccttca	aaccccccgtc	1500
cattcccttc	aggcaaaaca	gttttttggaa	cgctatttca	cgccgtggc	ggttgcaggc	1560
aacggaaagcc	tttgcgggtac	ggttacccgc	tattacgagc	cggtgtctaa	gggcgacgac	1620
aggcgacgg	cacaagcccg	cttcccgatt	tacggattt	ccgacgattt	tatctccgtc	1680
ccccctgcct	ccggttgcg	gagcgaaaa	gccctgtcc	gcatcaggca	gacggggaaaa	1740
aacagcggca	caatcgacaa	taccggcg	acacataccg	ccgaccttc	ccgattcccc	1800
atcacccgcgc	gcacaacggc	aatcaaaggc	aggtttgaag	gaagccgctt	cctccccc	1860
cacacgcgc	accaaaatcaa	cgccggcg	cttgcggca	aagcccgat	actcggttac	1920
gcccggaccc	ccgtcgaaact	ttttttatg	cacatccaa	gctggggcc	tctgaaaaaacc	1980
ccgtccggca	aatacatccg	catcggtat	gcccacaaaa	acgaacatcc	ctacgttcc	2040
atcgacgcct	atatggcgga	caaaggctac	ctcaagctcg	ggcagacactc	gtatcgaggc	2100
atcaaaggct	atatgcggca	aaatccgcaa	cgccctcgcc	aagtttggg	tcaaaacccc	2160
agctataatct	tttcccgcg	gcttgcgg	agcagcaatg	acggtcccgt	cggcgcactg	2220
ggcacgcgt	tgatggggga	atatgcggc	gcagtcgacc	ggcactacat	taccttgggc	2280
gcgccttat	tttgcggccac	cgcccatccg	gttacccgc	aagccctcaa	ccgcctgatt	2340
atggcgacgg	ataccggcg	cgcgattaaa	ggcgcgggtc	gcgtggatta	tttttgggg	2400
tagcgccacg	aaggccggcg	acttgcggc	aaacagaaaa	ccacgggtt	cgtctggcag	2460
ctccatccca	acqqtatqaa	qcccgaatac	qcggcgtaac	tcqag		2505

<210> 89  
<211> 83

<212> PRT  
<213> Artificial Sequence

<220>  
<223> deltaG287-919

<400> 89

Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala  
 1                   5                   10                   15  
 Ala Pro Val Val Ala Glu Lys Glu Thr Glu Val Lys Glu Asp Ala Pro

	20	25	30
5	Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Thr Gln Gly Ser Gln 35 35 40 45		
	Asp Met Ala Ala Val Ser Ala Glu Asn Thr Gly Asn Gly Gly Ala Ala 50 55 60		
10	Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Pro Gln Asn Asp Met 65 70 75 80		
	Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln Thr Gly Asn Asn Gln Pro 85 90 95		
15	Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser Asn Pro Ala Pro Ala Asn 100 105 110		
	Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala Asn Gly Val Leu Ile 115 120 125		
20	Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser 130 135 140		
	Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu Ala Pro Ser Lys Ser Glu 145 150 155 160		
	Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile Glu Lys Tyr Lys Lys Asp 165 170 175		
25	Gly Lys Ser Asp Lys Phe Thr Asn Leu Val Ala Thr Ala Val Gln Ala 180 185 190		
	Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr Lys Asp Lys Ser Ala Ser 195 200 205		
30	Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser Leu 210 215 220		
	Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu Ile 225 230 235 240		
35	Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile Phe 245 250 255		
	Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys Leu 260 265 270		
40	Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ala Lys Gly 275 280 285		
	Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His Phe 290 295 300		
45	His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg Gly Arg Phe Ala Ala 305 310 315 320		
	Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser Gly 325 330 335		
50	Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp Gly 340 345 350		
	Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Asp Val Ser 355 360 365		
55	Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly Lys Tyr Ser 370 375 380		
	Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val Phe Ala Gly		

## EP 1 790 660 A2

	385	390	395	400
5	Lys Lys Glu Gln Asp Gly Ser Gly Gly Gly Cys Gln Ser Lys Ser			
	405	410	415	
Ile Gln Thr Phe Pro Gln Pro Asp Thr Ser Val Ile Asn Gly Pro Asp				
	420	425	430	
10	Arg Pro Val Gly Ile Pro Asp Pro Ala Gly Thr Thr Val Gly Gly Gly			
	435	440	445	
Gly Ala Val Tyr Thr Val Val Pro His Leu Ser Leu Pro His Trp Ala				
	450	455	460	
15	Ala Gln Asp Phe Ala Lys Ser Leu Gln Ser Phe Arg Leu Gly Cys Ala			
	465	470	475	480
Asn Leu Lys Asn Arg Gln Gly Trp Gln Asp Val Cys Ala Gln Ala Phe				
	485	490	495	
Gln Thr Pro Val His Ser Phe Gln Ala Lys Gln Phe Phe Glu Arg Tyr				
	500	505	510	
20	Phe Thr Pro Trp Gln Val Ala Gly Asn Gly Ser Leu Ala Gly Thr Val			
	515	520	525	
Thr Gly Tyr Tyr Glu Pro Val Leu Lys Gly Asp Asp Arg Arg Thr Ala				
	530	535	540	
25	Gln Ala Arg Phe Pro Ile Tyr Gly Ile Pro Asp Asp Phe Ile Ser Val			
	545	550	555	560
Pro Leu Pro Ala Gly Leu Arg Ser Gly Lys Ala Leu Val Arg Ile Arg				
	565	570	575	
30	Gln Thr Gly Lys Asn Ser Gly Thr Ile Asp Asn Thr Gly Gly Thr His			
	580	585	590	
Thr Ala Asp Leu Ser Arg Phe Pro Ile Thr Ala Arg Thr Thr Ala Ile				
	595	600	605	
35	Lys Gly Arg Phe Glu Gly Ser Arg Phe Leu Pro Tyr His Thr Arg Asn			
	610	615	620	
Gln Ile Asn Gly Gly Ala Leu Asp Gly Lys Ala Pro Ile Leu Gly Tyr				
	625	630	635	640
40	Ala Glu Asp Pro Val Glu Leu Phe Phe Met His Ile Gln Gly Ser Gly			
	645	650	655	
Arg Leu Lys Thr Pro Ser Gly Lys Tyr Ile Arg Ile Gly Tyr Ala Asp				
	660	665	670	
45	Lys Asn Glu His Pro Tyr Val Ser Ile Gly Arg Tyr Met Ala Asp Lys			
	675	680	685	
Gly Tyr Leu Lys Leu Gly Gln Thr Ser Met Gln Gly Ile Lys Ala Tyr				
	690	695	700	
50	Met Arg Gln Asn Pro Gln Arg Leu Ala Glu Val Leu Gly Gln Asn Pro			
	705	710	715	720
Ser Tyr Ile Phe Phe Arg Glu Leu Ala Gly Ser Ser Asn Asp Gly Pro				
	725	730	735	
55	Val Gly Ala Leu Gly Thr Pro Leu Met Gly Glu Tyr Ala Gly Ala Val			
	740	745	750	
Asp Arg His Tyr Ile Thr Leu Gly Ala Pro Leu Phe Val Ala Thr Ala				

EP 1 790 660 A2

755                    760                    765

5	His Pro Val Thr Arg Lys Ala Leu Asn Arg Leu Ile Met Ala Gln Asp 770                    775                    780	
	Thr Gly Ser Ala Ile Lys Gly Ala Val Arg Val Asp Tyr Phe Trp Gly 785                    790                    795                    800	
10	Tyr Gly Asp Glu Ala Gly Glu Leu Ala Gly Lys Gln Lys Thr Thr Gly 805                    810                    815	
	Tyr Val Trp Gln Leu Leu Pro Asn Gly Met Lys Pro Glu Tyr Arg Pro 820                    825                    830	
15	<210> 90 <211> 1746 <212> DNA <213> Artificial Sequence	
20	<220> <223> deltaG287-953	
25	<400> 90 atggctagcc ccgatgttaa atcggccggac acgctgtcaa aaccggccgc tcctgttgc gctaaaaagg agacagaggt aaaagaagat gcgcacagg caggtttca aggacagg gcgcacatcca cacaaggcag ccaagatatg gcggcagttt cgccagaaaa tacaggca ggcggtgcgg caacaacggaa caaacccaaa aatgaagacg agggaccgca aaatgata ccgaaaatt ccgccgaatc cgcaaatcaa acagggaaaca accaaccgc cgattctt gatttgcggcc ccgcgtcaaa ccctgcaccc gcaatggcg gtagcaattt tggaaagggt gatttgcta atggcggtt gattgtatgg ccgtcgcaaa atataacgtt gaccactg aaaggcgatt ctgtaatgg tgataatttt ttggatggaa aagcaccgtc aaaaatcaga tttggaaaatt taatgtatgg tgaacgaaatt gagaatata agaaagatgg gaaaagcga aaatttacta atttgggtgc gacagcagtt caagctaatg gaactaacaa atatgtcat atttataaag acaaagtccgc ttcatcttca tctgcgcgtat tcaggcggtc tgcacgggt aggaggtcgc ttctcgccga gatggccgtat atccccgtca atcaggcgga tacgctgat gtcgatgggg aaggcggtcag cctgacgggg cattccggca atatcttcgc gcccgaagg aattaccggt atctgactta ggccggggaa aatttgcggc gggatcgta tgccctccg gtgcaaggcg aaccggcaaa aggcgaaatg ctgtctggca cggccgtgtc caacggcgaa gtgctgcatt ttcatacggaa aaacggccgt ccgtacccga ctagaggcag gtttgcgg aaagtgcatt tcggcagcaa atctgtggac ggcattatcg acagcggcga tgatttgc atgggtacgc aaaaattcaa agccgcattc gatggaaacg gcttaaggcg gacttggaa gaaaatggcg gggggatgt ttccggaaagg ttttacggcc cggccggcg ggaatgggg ggaaaataca gctatcgccc gacagatgcg gaaaaggcg gattcggtt gtttgcgg aaaaaagagc aggatggatc cggaggagga ggagccacct acaaagtggc cgaatatca gccaacgccc gtttcgcccc cgaccatttc aacaccagca ccaacgtcgg cgggtttt ggtctgaccg gtccgtcga gttcgaccaa gcaaaacgcg acggtaaat cgacatca atcccggtt ccaacctcga aagcggttcg caacacttta cgcaccaccc gaaatcag gacatcttcg atggccccc atatccggac atccgccttg ttccaccaa attcaactt aacggcaaaa aactggttt cgttgcggc aacctgacca tgacggcaa aaccggccca gtcaaactca aagccgaaaa attcaactgc taccaaagcc cgatggcgaa aaccgaagtg tgcggcggcg acttcagcac caccatcgac cgacccaaat gggcggttga ctacctcg aacgttggta tgacccaaag cgtccgcattt gacatccaaat tcgaggcagc caaacaata ctcgag	
30	<210> 91 <211> 579 <212> PRT <213> Artificial Sequence	
35	<220> <223> deltaG287-953	
40	<400> 91 Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala 1                    5                    10                    15	
45	Ala Pro Val Val Ala Glu Lys Glu Thr Glu Val Lys Glu Asp Ala Pro 20                    25                    30	

Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Thr Gln Gly Ser Gln  
 35 40 45  
 5 Asp Met Ala Ala Val Ser Ala Glu Asn Thr Gly Asn Gly Gly Ala Ala  
 50 55 60  
 Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Pro Gln Asn Asp Met  
 65 70 75 80  
 10 Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln Thr Gly Asn Asn Gln Pro  
 85 90 95  
 Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser Asn Pro Ala Pro Ala Asn  
 100 105 110  
 15 Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala Asn Gly Val Leu Ile  
 115 120 125  
 Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser  
 130 135 140  
 20 Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu Ala Pro Ser Lys Ser Glu  
 145 150 155 160  
 Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile Glu Lys Tyr Lys Lys Asp  
 165 170 175  
 25 Gly Lys Ser Asp Lys Phe Thr Asn Leu Val Ala Thr Ala Val Gln Ala  
 180 185 190  
 Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr Lys Asp Lys Ser Ala Ser  
 195 200 205  
 30 Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser Leu  
 210 215 220  
 Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu Ile  
 225 230 235 240  
 35 Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile Phe  
 245 250 255  
 Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys Leu  
 260 265 270  
 Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ala Lys Gly  
 275 280 285  
 40 Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His Phe  
 290 295 300  
 His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg Gly Arg Phe Ala Ala  
 305 310 315 320  
 45 Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser Gly  
 325 330 335  
 Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp Gly  
 340 345 350  
 50 Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Asp Val Ser  
 355 360 365  
 Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly Lys Tyr Ser  
 370 375 380  
 55 Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val Phe Ala Gly  
 385 390 395 400

Lys Lys Glu Gln Asp Gly Ser Gly Gly Gly Ala Thr Tyr Lys Val  
 405 410 415

5 Asp Glu Tyr His Ala Asn Ala Arg Phe Ala Ile Asp His Phe Asn Thr  
 420 425 430

Ser Thr Asn Val Gly Gly Phe Tyr Gly Leu Thr Gly Ser Val Glu Phe  
 435 440 445

10 Asp Gln Ala Lys Arg Asp Gly Lys Ile Asp Ile Thr Ile Pro Val Ala  
 450 455 460

Asn Leu Gln Ser Gly Ser Gln His Phe Thr Asp His Leu Lys Ser Ala  
 465 470 475 480

15 Asp Ile Phe Asp Ala Ala Gln Tyr Pro Asp Ile Arg Phe Val Ser Thr  
 485 490 495

Lys Phe Asn Phe Asn Gly Lys Lys Leu Val Ser Val Asp Gly Asn Leu  
 500 505 510

20 Thr Met His Gly Lys Thr Ala Pro Val Lys Leu Lys Ala Glu Lys Phe  
 515 520 525

Asn Cys Tyr Gln Ser Pro Met Ala Lys Thr Glu Val Cys Gly Gly Asp  
 530 535 540

25 Phe Ser Thr Thr Ile Asp Arg Thr Lys Trp Gly Val Asp Tyr Leu Val  
 545 550 555 560

Asn Val Gly Met Thr Lys Ser Val Arg Ile Asp Ile Gln Ile Glu Ala  
 565 570 575

Ala Lys Gln

30 <210> 92  
 <211> 2388  
 <212> DNA  
 <213> Artificial Sequence

35 <220>  
 <223> deltaG287-961

40 <400> 92  
 atggctagcc ccgatgttaa atcggcgac acgctgtcaa aaccggccgc tcctgttgtt 60  
 gctaaaaag agacagaggt aaaagaagat gcgccacagg caggttctca aggacaggc 120  
 gcccattcca cacaaggcag ccaagatatg gcggcagttt cgccagaaaa tacaggcaat 180  
 ggcggtcgg caacaacgga caaacccaaa aatgaagacg agggaccgca aaatgatatg 240  
 cgc当地 att ccgcgaatc cgc当地 acagggaaaca accaaccgcg cgattctca 300  
 gat tccgccc ccgcgtcaaa ccctgcacct gc当地ggcg gt当地caattt tgaaagggtt 360  
 gat tggctta atggctttt gattgatggg cc当地c当地aaa atataacgtt gaccactgt 420  
 aaaggcgatt ct当地atgg tt当地atttt tt当地atggaa aagaccgctc aaaatcagaa 480  
 tttgaaaatt taaaatgatc tgaacgatt gagaatata agaaagatgg gaaaagcgat 540  
 aaatttacta atttgggtgc gacagcattt caagctaattt gaactaaca atatgtcatc 600  
 atttataaag acaagtccgc ttcatcttca tctgc当地atc tcaggcgttc tgacggctg 660  
 aggaggcgc ttc当地gccc当地 gatccgctt当地 atccccgtca atcaggc当地 tacgctgatt 720  
 gtc当地atggg aaggcggttag cctgacgggg catttccggca atatccgc当地 gccc当地agggg 780  
 aatttccggt atctgactt当地 cggggcgaa aaatttccggca gcgatcgta tgccctccgt 840  
 50 gtc当地aggcg aaccggcaaa aggcaatg cttgctggca cggccgtt当地 caacggcgaa 900  
 gtc当地gtcatt tt当地atcgaa aaacggccgt cc当地tccggca ctagaggc当地 gtttccgc当地 960  
 aaaggcgatt tcggc当地caaa atctgtggac ggc当地tccggca acaggcgca tgat tttgcat 1020  
 atgggtacgc aaaaatttcaaa agccgccatc gatggaaacg gctttaagggg gacttggacg 1080  
 gaaaatggcg gccc当地ggatgt ttccggaaagg tttacggcc cggccggc当地 ggaatggcg 1140  
 ggaaaatataa gctatcgccc gacagatcg gaaaaggcg gattcggc当地 gtttccggc当地 1200  
 55 aaaaatggagc aggatggatc cggaggagga ggagccacaa acgacgacgaa tgat taaa 1260  
 gctggagaga ccatctacga cattgatgaa gacggc当地aa aagaaatcaa cggttccaa 1320  
 gctggagaga ccatctacga cattgatgaa gacggc当地aa aagaaatcaa cggttccaa 1380

	gcagccgatg ttgaagccga cgactttaaa ggtctgggtc taaaaaaagt cgtgactaac	1440
	ctgaccaaaa ccgtcaatga aaacaaacaa aacgtcgatg ccaaaggtaaa agctgcagaa	1500
5	tctgaaatag aaaagttaac aaccaagttt gcagacactg atgcccgttt agcagatact	1560
	gatgccgcgc tggatgcac caccAACGCC ttgataaat tgggagaaaaataaAcgaca	1620
	tttgctgaag agactaagac aaatatcgtaaaaattgtt aaaaattaga agccgtggct	1680
	gataccgtcg acaagcatgc cgaagcattt aacgatatcg ccgattcatt ggatgaaacc	1740
	aacactaagg cagacgaagc cgtcaaaacc gccaatgaag ccaaAcagac gcccgaagaa	1800
	acccaaacaaa acgtcgatgc caaagtaaaa gctgcagaaaa ctgcagcagg caaAgccgaa	1860
10	gctgccgcgtc gcacagctaa tactcgaccc gacaaggccg aagctgtcgc tgcaaaagtt	1920
	accgacatca aagctgatcatcgtaaaagataata ttgctaaaaaa agcaaAcagt	1980
	gcccgtcg acaccagaga agagtctgac agcaaatttg tcagaatttg tggtctgaac	2040
	gctactaccg aaaaatttggc cacacgctt gcttcgtc aaaaatccat tgccgatcac	2100
	gatactcgcc tgaacggttt ggataaaaaca gtgtcagacc tgccaaaga aaccggccaa	2160
	ggccttgcag aacaagccgc gctctccgtt ctgttccaac cttacaacgt gggtcggttc	2220
	aatgttaacgg ctgcagtccgg cggctacaaa tccgaatcgg cagtcgccccat cggtaCCGc	2280
	ttccgctttt ccgaaaactt tgccgccaa gcagcgtgg cagtcggcac ttcgctccgg	2340
15	tcttccgcag cctaccatgt cggcgtcaat tacgagtggt aactcgag	2388
	<210> 93	
	<211> 793	
	<212> PRT	
	<213> Artificial Sequence	
20	<220>	
	<223> deltaG287-961	
	<400> 93	
	Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala	
	1 5 10 15	
25	Ala Pro Val Val Ala Glu Lys Glu Thr Glu Val Lys Glu Asp Ala Pro	
	20 25 30	
	Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Thr Gln Gly Ser Gln	
	35 40 45	
30	Asp Met Ala Ala Val Ser Ala Glu Asn Thr Gly Asn Gly Gly Ala Ala	
	50 55 60	
	Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Pro Gln Asn Asp Met	
	65 70 75 80	
35	Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln Thr Gly Asn Asn Gln Pro	
	85 90 95	
	Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser Asn Pro Ala Pro Ala Asn	
	100 105 110	
40	Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala Asn Gly Val Leu Ile	
	115 120 125	
	Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser	
	130 135 140	
45	Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu Ala Pro Ser Lys Ser Glu	
	145 150 155 160	
	Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile Glu Lys Tyr Lys Lys Asp	
	165 170 175	
50	Gly Lys Ser Asp Lys Phe Thr Asn Leu Val Ala Thr Ala Val Gln Ala	
	180 185 190	
	Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr Lys Asp Lys Ser Ala Ser	
	195 200 205	
55	Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser Leu	
	210 215 220	

Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu Ile  
 225 230 235 240  
 5 Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile Phe  
 245 250 255  
 Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys Leu  
 260 265 270  
 10 Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ala Lys Gly  
 275 280 285  
 Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His Phe  
 290 295 300  
 15 His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg Gly Arg Phe Ala Ala  
 305 310 315 320  
 Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser Gly  
 325 330 335  
 20 Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp Gly  
 340 345 350  
 Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Asp Val Ser  
 355 360 365  
 25 Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly Lys Tyr Ser  
 370 375 380  
 Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val Phe Ala Gly  
 385 390 395 400  
 Lys Lys Glu Gln Asp Gly Ser Gly Gly Ala Thr Asn Asp Asp  
 405 410 415  
 30 Asp Val Lys Lys Ala Ala Thr Val Ala Ile Ala Ala Ala Tyr Asn Asn  
 420 425 430  
 Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile  
 435 440 445  
 35 Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp Ala Thr Ala Ala Asp Val  
 450 455 460  
 Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu Lys Lys Val Val Thr Asn  
 465 470 475 480  
 40 Leu Thr Lys Thr Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val  
 485 490 495  
 Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp  
 500 505 510  
 45 Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala Ala Leu Asp Ala Thr Thr  
 515 520 525  
 Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu  
 530 535 540  
 50 Thr Lys Thr Asn Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala  
 545 550 555 560  
 Asp Thr Val Asp Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp Ser  
 565 570 575  
 55 Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala Asn  
 580 585 590

## EP 1 790 660 A2

Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala Lys  
 595 600 605  
 5 Val Lys Ala Ala Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Ala Gly  
 610 615 620  
 Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys Val  
 625 630 635 640  
 10 Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn Lys Asp Asn Ile Ala Lys  
 645 650 655  
 Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg Glu Glu Ser Asp Ser Lys  
 660 665 670  
 Phe Val Arg Ile Asp Gly Leu Asn Ala Thr Thr Glu Lys Leu Asp Thr  
 15 675 680 685  
 Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala Asp His Asp Thr Arg Leu  
 690 695 700  
 20 Asn Gly Leu Asp Lys Thr Val Ser Asp Leu Arg Lys Glu Thr Arg Gln  
 705 710 715 720  
 Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu Phe Gln Pro Tyr Asn  
 725 730 735  
 25 Val Gly Arg Phe Asn Val Thr Ala Ala Val Gly Gly Tyr Lys Ser Glu  
 740 745 750  
 Ser Ala Val Ala Ile Gly Thr Gly Phe Arg Phe Thr Glu Asn Phe Ala  
 755 760 765  
 30 Ala Lys Ala Gly Val Ala Val Gly Thr Ser Ser Gly Ser Ser Ala Ala  
 770 775 780  
 Tyr His Val Gly Val Asn Tyr Glu Trp  
 785 790  
 35 <210> 94  
 <211> 2700  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> deltaG287NZ-919  
 40 <400> 94  
 atggctagcc ccgatgtcaa gtcggcgac acgcgtcaa aacctgccgc ccctgttgg  
 tctaaaaag agacagaggc aaaggaagat gcgcacagg caggttctca aggacaggc  
 60 gcgccatccg cacaaggcg cacaaggcg tcaagatatg gcggcggtt cggaaagaaa tacaggcaat  
 ggcggtgccg cagcaacgg aaaaaaaaaa aatgaagacg agggggcgca aaatgatatg  
 120 cccaaaaatg cccggatac agatgttt acacccaatc acaccccggc ttcaaatatg  
 cccggcgaa atatggaaaaa ccaaggaccg gatgggggg aatcgagac gccggcaa  
 180 caaccggata tggcaatac ggcggacggg atgcagggtg acgatccgtc ggcaggcgg  
 gaaaatgccg gcaatacgcc tgcccaaggt acaaatcaag ccgaaaacaa tcaaaccgccc  
 240 gttctcaaa atcctgcctc ttcaaccaat cctagcgcc agaardcg tggatgttt  
 ggaaggacga acgtggggca ttctgttgg attgacgggc cgatcgaaaaataacgttg  
 300 acccaactgt aaggcgattt ttgtatgtgc aataattttct tggatgttgg agtacagcta  
 360 aataatcgaaat ttggaaaaatt aagtgtatgc gacaaaaataa gtaattacaa gaaagatggg  
 420 aagaatgacg ggaagaatgaa taaatttgc gtttgggtt ccgatgttgt gcatgttgc  
 480 ggaatcaatc aataatattat cttttataaa cctaaacccca cttcatttgc gcatgtttagg  
 540 cgttctgcac ggtcgaggcg gtcgcttccg gccgagatgc cgatcgatcc cgtcaatcag  
 600 gcgatcgatcg tgatgtgc tgggggaaacg gtcagccgtc cggggcattc cggcaatatc  
 660 ttcgcgcggc aaggaaatta ccggatcgatcg acttacgggg cggaaaaattt gccccggcgaa  
 720 tcgtatgccc tccgttgc aggcgaaacct tcaaaaggcg aaatgtcgc gggcacggca  
 780 gtgtacaacg gcaagtgct gcatttcat acggaaaaacg gccgtccgtc cccgtccaga  
 840 ggcaggttt ccgcaaaagt cgatccggc agcaatctg tggacggcat tatcgacagc  
 900 ggcgatggtt tgcatatggg tacgcaaaaa ttcaaagccg ccatcgatgg aaacggctt  
 960 1020 1080 1140 1200 1260

	aaggggactt	ggacggaaaa	tggcggcggt	gatgtttccg	gaaagttta	cggccgggcc	1320
5	ggcgaggaaag	tggcggaaaa	atacagctat	cgcccaacag	atgcggaaa	gggcggattc	1380
	ggcgttttgc	ccggcaaaaa	agagcaggat	ggatccggag	gaggaggatg	ccaaagcaag	1440
	agcatccaaa	ccttccgca	acccgacaca	tccgtcatca	acggccccgg	ccggccggtc	1500
	ggcatccccg	accccggccg	aacgacggtc	ggccggcgcc	gggcccgtct	taccgttgta	1560
	ccgcacctgt	ccctggccca	ctggcggcg	caggatttcg	caaaaaggct	gcaatccctc	1620
	cgcctcgct	gcccattt	gaaaaaacgc	caaggctggc	aggatgtgt	cgcctaagcc	1680
10	tttcaaacc	ccgtccattc	ctttcaggca	aaacagtttt	ttgaacgctt	tttcacggcc	1740
	tggcaggtt	caggcaacgg	aagccttgcc	ggtacggta	ccggctatta	cgagccggtg	1800
	ctgaaggcg	acgacaggcg	gacggcacaa	gcccgttcc	cgatttacgg	tattcccgac	1860
	gattttatct	ccgtccccct	gcctggccgt	ttgcggagcg	aaaaagccct	tgtccgcattc	1920
	aggcagacg	gaaaaaacag	cggcacaatc	gacaataccg	ggggcacaca	taccggccac	1980
15	ctctcccgat	tcccccattac	cgcgcgca	acggcaatca	aaggcagggt	tgaaggaagc	2040
	cgttccctcc	ccttacacac	gwgcaaccaa	atcaacggcg	gwgcttga	cggcaaagcc	2100
	ccgatactcg	gttacggcg	agaccccg	gaacttttt	ttatgcacat	ccaaggctcg	2160
	ggccgtctga	aaaccccg	cgwgcaatac	atccgcatcg	gctatgcca	aaaaaacgaa	2220
	catccctact	ttttccatcg	acgctatatg	gcggacaaag	gctacctcaa	gctccggcag	2280
	acctcgatgc	aggcgtca	acgctatatg	cgwgcaaaatc	cgcaacgct	cgccgaagtt	2340
	ttgggtcaaa	accccgacta	tatcttttc	cgcgagctt	cgwgagcag	aatgacggt	2400
	cccgccggcg	cactgggcac	gwggttgc	ggggatatg	cgwgccgagt	cgacccggc	2460
20	tacattac	ttggcgcg	cttatttgc	gwgaccggcc	atccggttac	ccgcaaagcc	2520
	ctcaaccg	tgattatggc	gcaggatacc	ggcagcgcga	ttaaaggcgc	ggtgcgcgt	2580
	gatttttt	ggggatacgg	cgacgaagcc	ggcgaactt	ccggcaaaca	aaaaaccacg	2640
	ggttacgtct	ggcagctct	acccaacgg	atgaagcc	aataccggcc	gtaaaagctt	2700
25	<210>	95					
	<211>	897					
	<212>	PRT					
	<213>	Artificial Sequence					
	<220>						
	<223>	deltaG287NZ-919					
30	<400>	95					
	Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala						
	1	5	10		15		
	Ala Pro Val Val Ser Glu Lys Glu Thr Glu Ala Lys Glu Asp Ala Pro						
	20	25	30				
35	Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Ala Gln Gly Gly Gln						
	35	40	45				
	Asp Met Ala Ala Val Ser Glu Glu Asn Thr Gly Asn Gly Gly Ala Ala						
	50	55	60				
40	Ala Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Ala Gln Asn Asp Met						
	65	70	75	80			
	Pro Gln Asn Ala Ala Asp Thr Asp Ser Leu Thr Pro Asn His Thr Pro						
	85	90	95				
45	Ala Ser Asn Met Pro Ala Gly Asn Met Glu Asn Gln Ala Pro Asp Ala						
	100	105	110				
	Gly Glu Ser Glu Gln Pro Ala Asn Gln Pro Asp Met Ala Asn Thr Ala						
	115	120	125				
	Asp Gly Met Gln Gly Asp Asp Pro Ser Ala Gly Gly Glu Asn Ala Gly						
	130	135	140				
50	Asn Thr Ala Ala Gln Gly Thr Asn Gln Ala Glu Asn Asn Gln Thr Ala						
	145	150	155	160			
	Gly Ser Gln Asn Pro Ala Ser Ser Thr Asn Pro Ser Ala Thr Asn Ser						
	165	170	175				
55	Gly Gly Asp Phe Gly Arg Thr Asn Val Gly Asn Ser Val Val Ile Asp						
	180	185	190				

Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser Cys  
 195 200 205  
 5 Ser Gly Asn Asn Phe Leu Asp Glu Glu Val Gln Leu Lys Ser Glu Phe  
 210 215 220  
 Glu Lys Leu Ser Asp Ala Asp Lys Ile Ser Asn Tyr Lys Lys Asp Gly  
 225 230 235 240  
 10 Lys Asn Asp Gly Lys Asn Asp Lys Phe Val Gly Leu Val Ala Asp Ser  
 245 250 255  
 Val Gln Met Lys Gly Ile Asn Gln Tyr Ile Ile Phe Tyr Lys Pro Lys  
 260 265 270  
 15 Pro Thr Ser Phe Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser  
 275 280 285  
 Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu  
 290 295 300  
 20 Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile  
 305 310 315 320  
 Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys  
 325 330 335  
 25 Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ser Lys  
 340 345 350  
 Gly Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His  
 355 360 365  
 30 Phe His Thr Glu Asn Gly Arg Pro Ser Pro Ser Arg Gly Arg Phe Ala  
 370 375 380  
 Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser  
 385 390 395 400  
 35 Gly Asp Gly Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp  
 405 410 415  
 Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Asp Val  
 420 425 430  
 40 Ser Gly Lys Phe Tyr Gly Pro Ala Gly Glu Val Ala Gly Lys Tyr  
 435 440 445  
 Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Phe Gly Val Phe Ala  
 450 455 460  
 Gly Lys Lys Glu Gln Asp Gly Ser Gly Gly Gly Cys Gln Ser Lys  
 465 470 475 480  
 45 Ser Ile Gln Thr Phe Pro Gln Pro Asp Thr Ser Val Ile Asn Gly Pro  
 485 490 495  
 Asp Arg Pro Val Gly Ile Pro Asp Pro Ala Gly Thr Thr Val Gly Gly  
 500 505 510  
 50 Gly Gly Ala Val Tyr Thr Val Val Pro His Leu Ser Leu Pro His Trp  
 515 520 525  
 Ala Ala Gln Asp Phe Ala Lys Ser Leu Gln Ser Phe Arg Leu Gly Cys  
 530 535 540  
 55 Ala Asn Leu Lys Asn Arg Gln Gly Trp Gln Asp Val Cys Ala Gln Ala  
 545 550 555 560

## EP 1 790 660 A2

Phe Gln Thr Pro Val His Ser Phe Gln Ala Lys Gln Phe Phe Glu Arg  
 565 570 575  
 5 Tyr Phe Thr Pro Trp Gln Val Ala Gly Asn Gly Ser Leu Ala Gly Thr  
 580 585 590  
 Val Thr Gly Tyr Tyr Glu Pro Val Leu Lys Gly Asp Asp Arg Arg Thr  
 595 600 605  
 10 Ala Gln Ala Arg Phe Pro Ile Tyr Gly Ile Pro Asp Asp Phe Ile Ser  
 610 615 620  
 Val Pro Leu Pro Ala Gly Leu Arg Ser Gly Lys Ala Leu Val Arg Ile  
 625 630 635 640  
 15 Arg Gln Thr Gly Lys Asn Ser Gly Thr Ile Asp Asn Thr Gly Gly Thr  
 645 650 655  
 His Thr Ala Asp Leu Ser Arg Phe Pro Ile Thr Ala Arg Thr Thr Ala  
 660 665 670  
 20 Ile Lys Gly Arg Phe Glu Gly Ser Arg Phe Leu Pro Tyr His Thr Arg  
 675 680 685  
 Asn Gln Ile Asn Gly Gly Ala Leu Asp Gly Lys Ala Pro Ile Leu Gly  
 690 695 700  
 25 Tyr Ala Glu Asp Pro Val Glu Leu Phe Phe Met His Ile Gln Gly Ser  
 705 710 715 720  
 Gly Arg Leu Lys Thr Pro Ser Gly Lys Tyr Ile Arg Ile Gly Tyr Ala  
 725 730 735  
 30 Asp Lys Asn Glu His Pro Tyr Val Ser Ile Gly Arg Tyr Met Ala Asp  
 740 745 750  
 Lys Gly Tyr Leu Lys Leu Gly Gln Thr Ser Met Gln Gly Ile Lys Ala  
 755 760 765  
 35 Tyr Met Arg Gln Asn Pro Gln Arg Leu Ala Glu Val Leu Gly Gln Asn  
 770 775 780  
 Pro Ser Tyr Ile Phe Phe Arg Glu Leu Ala Gly Ser Ser Asn Asp Gly  
 785 790 795 800  
 40 Pro Val Gly Ala Leu Gly Thr Pro Leu Met Gly Glu Tyr Ala Gly Ala  
 805 810 815  
 Val Asp Arg His Tyr Ile Thr Leu Gly Ala Pro Leu Phe Val Ala Thr  
 820 825 830  
 45 Ala His Pro Val Thr Arg Lys Ala Leu Asn Arg Leu Ile Met Ala Gln  
 835 840 845  
 Asp Thr Gly Ser Ala Ile Lys Gly Ala Val Arg Val Asp Tyr Phe Trp  
 850 855 860  
 Gly Tyr Gly Asp Glu Ala Gly Glu Leu Ala Gly Lys Gln Lys Thr Thr  
 865 870 875 880  
 50 Gly Tyr Val Trp Gln Leu Leu Pro Asn Gly Met Lys Pro Glu Tyr Arg  
 885 890 895  
 Pro  
 55 <210> 96  
 <211> 1941

**<212> DNA**  
**<213> Artificial Sequence**

**5 <220>**  
**<223> deltaG287NZ-953**

<b>&lt;400&gt; 96</b> atggctagcc ccgatgtcaa gtcggcggac acgctgtcaa aacctgccgc ccctgttgtt tctgaaaaag agacagaggc aaaggaagat gcgccacagg caggttctca aggacaggc gcccacatccg cacaaggcgg tcaagatatg ggcggcggtt cggaaagaaaa tacaggcaat ggcgggtcgcc cagaatccgg aaaaacccaa aatgaagacg agggggcga aaatgatatg ccgcaaatatg ccgcccatac agataggta acaccgaatc acaccccgcc ttcaaatatg ccggccggaa atatggaaaa ccaagcaccg gatgccgggg aatcgagca gcccggaaac caaccggata tgcaaaatac ggcggacggg atgcagggtg acgatccgtc ggcaggcggg gaaaatggcg gcaatacggc tgcccaaggt acaaatacaag cggaaaacaa tcaaaccgcc ggttctcaaa atccctgcctc ttcaaccaat cctagcgcca cgaatagcgg tggtgatttt ggaaggacga acgtgggcaat ttctgtgtg attgacgggc cgtcgcaaaa tataacgttg acccactgt aaggcgattc ttgtatgtc aataatttct tggatgaga agtacagcta aaatcagaat ttgaaaaatt aagtgtatca gacaaaaataa gtaattacaa gaaagatggg aagaatgacg ggaagaatga taaatttgc ggttgggtt cgcgatgtgt gcagatgaag ggaatcaatc aatatattat cttaataaaa cctaaaccca cttcatttgc gcgatttagg cgttctgcac ggtcgaggcg gtcgctccg gccgagatgc cgctgattcc cgtaatcag gcccggatcc tgattgtcg tggggaaagc gtcagccgtc cggggcatcc cggcaatatac 20 ttcgcgcccc aagggaatta cccgttatcg acttacgggg cggaaaatt gcccggcga tcgtatgtccc tcgcgttca aggcaacatc tcaaaaaggcg aaatgctcgc gggcacggca gtgtacaacg gcgaaagtgtc gcatttcat acgaaaaacg gccgtccgtc cccgtccaga ggcagggttgc cgcggaaatgt cgatttgcg agcaaatctg tggacggcat tatcgacagc ggcgtatgtt tgcataatggg tacgaaaaaa ttcaaaagccg ccatcgatgg aaacggctt aaggggactt ggacggaaaaa tggcggcggg gatgtttccg gaaagttt a cggcccgcc ggcgaggaaag tggcggggaaa atacagctat cgcggaaacag atgcggaaaaa gggcgattc ggcgttgcggg cggcaaaaaa agacaggat ggatccggag gaggaggagc cacttacaaa gtggacgaat atcacgccaa cggccgttgc gccatcgacc atttcaacac cagcaccaac gtcggcggtt ttacgggtc gaccgggtcc gtcgagttcg accaagcaaa acgcgacggt aaaatcgaca tcaccatccc cgttgcac ctcgaaagcg gttcgcaaca ctttaccgac cacctgaaaat cagccgacat ctgcgtatgc gcccaatatac cggacatccg ctttgtttcc accaaaattca acttcaacgg caaaaatcg gttccgttgc acggcaacatc gaccatgcac ggcaaaaacccg ccccccgtcaa actcaaaagcc gaaaaattca actgctacca aagcccgatg gcaaaaacccg aagtttgcgg cggcgacttc agcaccacca tcgaccgcac caaatggggc gtggacttacc tcgttaacgt tggtatgacc aaaagcgtcc gcatcgacat ccaaatacgag gcagccaaac aataaaaagct t	60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620 1680 1740 1800 1860 1920 1941
<b>35 &lt;210&gt; 97</b> <b>&lt;211&gt; 644</b> <b>&lt;212&gt; PRT</b> <b>&lt;213&gt; Artificial Sequence</b>	
<b>40 &lt;220&gt;</b> <b>&lt;223&gt; deltaG287NZ-953</b>	
<b>&lt;400&gt; 97</b> Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala 1 5 10 15	Ala Pro Val Val Ser Glu Lys Glu Thr Glu Ala Lys Glu Asp Ala Pro 20 25 30
Ala Pro Val Val Ser Glu Lys Glu Thr Glu Ala Lys Glu Asp Ala Pro 35 40 45	Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Ala Gln Gly Gly Gln 50 55 60
Asp Met Ala Ala Val Ser Glu Glu Asn Thr Gly Asn Gly Gly Ala Ala 65 70 75	Ala Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Ala Gln Asn Asp Met 85 90 95
Pro Gln Asn Ala Ala Asp Thr Asp Ser Leu Thr Pro Asn His Thr Pro 100 105 110	Ala Ser Asn Met Pro Ala Gly Asn Met Glu Asn Gln Ala Pro Asp Ala 115 120 125

## EP 1 790 660 A2

	100	105	110
5	Gly Glu Ser Glu Gln Pro Ala Asn Gln Pro Asp Met Ala Asn Thr Ala 115 120 125		
	Asp Gly Met Gln Gly Asp Asp Pro Ser Ala Gly Gly Glu Asn Ala Gly 130 135 140		
10	Asn Thr Ala Ala Gln Gly Thr Asn Gln Ala Glu Asn Asn Gln Thr Ala 145 150 155 160		
	Gly Ser Gln Asn Pro Ala Ser Ser Thr Asn Pro Ser Ala Thr Asn Ser 165 170 175		
15	Gly Gly Asp Phe Gly Arg Thr Asn Val Gly Asn Ser Val Val Ile Asp 180 185 190		
	Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser Cys 195 200 205		
	Ser Gly Asn Asn Phe Leu Asp Glu Glu Val Gln Leu Lys Ser Glu Phe 210 215 220		
20	Glu Lys Leu Ser Asp Ala Asp Lys Ile Ser Asn Tyr Lys Lys Asp Gly 225 230 235 240		
	Lys Asn Asp Gly Lys Asn Asp Lys Phe Val Gly Leu Val Ala Asp Ser 245 250 255		
25	Val Gln Met Lys Gly Ile Asn Gln Tyr Ile Ile Phe Tyr Lys Pro Lys 260 265 270		
	Pro Thr Ser Phe Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser 275 280 285		
30	Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu 290 295 300		
	Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile 305 310 315 320		
35	Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys 325 330 335		
	Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ser Lys 340 345 350		
40	Gly Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His 355 360 365		
	Phe His Thr Glu Asn Gly Arg Pro Ser Pro Ser Arg Gly Arg Phe Ala 370 375 380		
45	Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser 385 390 395 400		
	Gly Asp Gly Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp 405 410 415		
50	Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Asp Val 420 425 430		
	Ser Gly Lys Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly Lys Tyr 435 440 445		
55	Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val Phe Ala 450 455 460		
	Gly Lys Lys Glu Gln Asp Gly Ser Gly Gly Ala Thr Tyr Lys		

465	470	475	480
	Val Asp Glu Tyr His Ala Asn Ala Arg Phe Ala Ile Asp His Phe Asn		
5	485	490	495
	Thr Ser Thr Asn Val Gly Gly Phe Tyr Gly Leu Thr Gly Ser Val Glu		
	500	505	510
	Phe Asp Gln Ala Lys Arg Asp Gly Lys Ile Asp Ile Thr Ile Pro Val		
10	515	520	525
	Ala Asn Leu Gln Ser Gly Ser Gln His Phe Thr Asp His Leu Lys Ser		
	530	535	540
	Ala Asp Ile Phe Asp Ala Ala Gln Tyr Pro Asp Ile Arg Phe Val Ser		
15	545	550	555
	Thr Lys Phe Asn Phe Asn Gly Lys Lys Leu Val Ser Val Asp Gly Asn		
	565	570	575
	Leu Thr Met His Gly Lys Thr Ala Pro Val Lys Leu Lys Ala Glu Lys		
20	580	585	590
	Phe Asn Cys Tyr Gln Ser Pro Met Ala Lys Thr Glu Val Cys Gly Gly		
	595	600	605
	Asp Phe Ser Thr Thr Ile Asp Arg Thr Lys Trp Gly Val Asp Tyr Leu		
25	610	615	620
	Val Asn Val Gly Met Thr Lys Ser Val Arg Ile Asp Ile Gln Ile Glu		
	625	630	635
	Ala Ala Lys Gln		
30	<210> 98		
	<211> 2583		
	<212> DNA		
	<213> Artificial Sequence		
35	<220>		
	<223> deltaG287NZ-961		
	<400> 98		
	atggctagcc ccgatgtcaa gtcggccgac acgctgtcaa aacctggccgc ccctgttgtt		60
	tctgaaaaag agacagaggc aaaggaagat gcgccacagg caggttctca aggacaggc		120
	gcccattcg cacaaggcgg tcaagatatg gcggcggttt cggaaagaaaa tacaggcaat		180
	ggcggtcgg cagcaacggg caaaccggaa atatggaaatc agggggcgca aaatgtatgg		240
	ccgaaaatg cccggataatc agatagggtt acacggaaatc acacccggc ttcaaatatg		300
	ccggccggaa atatggaaaaa ccaagcaccg gatgccccggg aatcgagca gccggcaaac		360
	caaccggata tggcaaaatac ggcggacggg atgcagggtt acgatccgtc ggcaggcggg		420
	gaaaatgcgg gcaatacggc tgcccaaggt acaaataatc aatcgagcg tggtgatttt		480
	ggttctcaaa atccctgcctc ttcaaccat cctagcgccca cgaatagcg tggtgatttt		540
	ggaaggacga acgtggggaa ttctgttgtt attgacggggc cgtcgaaaaa tataacgttg		600
	accacactgtt aaggcgattt ttgtgtggc aataatttttctt tggatgaaga agtacagcta		660
	aaatcagaat ttggaaaaattt aagtgtatgc gacaaaataaa gtaattacaa gaaagatggg		720
	aagaatgcg ggaagaatga taaattttgtc gttttgttgc ccatagatgt gcagatgaag		780
	ggaatcaatc aatatattat cttttataaa cttaaacccca cttcatttgc gcgatattagg		840
	cgttctgcac ggtcgaggcg gtcgctccg gccgagatgc cgctgatccc cgtaatcag		900
	cgggatacgc tgattgtcg tggggaaatcg gtacgttgc cggggatttgc cggcaatatc		960
	ttcgcgcccc aagggaattt ccggtatctg acttacgggg cggaaaaattt gcccggcgga		1020
	tcgtatgccc tcctgttca aggcaaccc tcaaaaggcg aatgtctgc gggcacggca		1080
	gtgtacaacg gcaaggatgtc gcattttcat acggaaaacg gccgtccgtc cccgtccaga		1140
	ggcagggttg ccgcggaaatcg cgatttcggc agcaaatctg tggacggcat tatcgacagc		1200
	ggcgtatgggtt tgcataatggg tacgcggaaa ttcaaaatcg ccatcgatgg aaacggcttt		1260
	aaggggactt ggacggaaaaa tggcgccggg gatgtttcg gaaatgttta cggcccgccc		1320
	ggcgaggaaatggcgaaaaa atacagctat cggccaaacag atgcggaaaaa gggcgatttgc		1380
	ggcgtgtttt ccggcaaaaaa agagcaggat ggtatccggag gaggaggagc cacaacgcac		1440
	gacgatgttta aaaaagctgc cactgtggcc attgtctgtc cctacaacaa tggccaagaa		1500

	atcaacgggtt tcaaaggctgg agagaccatc tacgacattg atgaagacgg cacaattacc	1560
	aaaaaaagacg caactgcagc cgatgttcaa gcccacgact ttaaaggctct gggctctgaaa	1620
5	aaagtctgtga ctaacacctac caaaaacccgtc aatgaaaaca aacaaaacgt cgatgccaaa	1680
	gtaaaaggctg cagaatctga aatagaaaag ttaacaaccca agtttagcaga cactgatgcc	1740
	gcttttagcag atactgtatcg cgctctggat gcaaccacca acgccttggaa taaattggga	1800
	gaaaatataa cgacatttgcc tgaagagact aagacaata tcgtaaaaat tgatgaaaaaa	1860
	ttagaagccg tggctgatac cgtcgacaaag catgcccgaag cattcaacga tatcgccgat	1920
10	tcattggatg aaacccaacac taaggcagac gaagccgtca aaaccgc当地 tgaagccaaa	1980
	cagacggccg aagaaaccaa acaaaaacgtc gatgccaaag taaaagctgc agaaaactgca	2040
	gcaggccaaag ccgaagctgc cgctggcaca gctaataactg cagccgacaa ggccgaaagct	2100
	gtcgctgcaa aagttaaccga catcaaaagct gatatcgcta cgaacaaaga taatattgct	2160
	aaaaaaagacg acatgtccgc cgtgtacacc agagaagagt ctgacagc当地 attttgtcaga	2220
	attgtatggc tgaacgctac taccggaaaaa ttggacacac gcttggcttc tgctgaaaaaa	2280
15	tccattggcg atcacgatac tcgcctgtaaac ggtttggata aaacagtgtc agaccctgc当地	2340
	aaagaaaccc gccaaggcct tgcaaaacaa gcccgc当地 ccggctgtt ccaaccttac	2400
	aacgtgggtc ggttcaatgt aacggctgca gtccggcggct acaaatccga atcggcagtc	2460
	gccatcggtt ccggcttccg ctttaccgaa aacttgc当地 ccaaggcagg cgtggcagtc	2520
	ggcacttcgt ccggcttccg cgcagctac catgtcggc当地 tcaattacga gtggtaaaag	2580
	ctt	2583
20	<210> 99 <211> 858 <212> PRT <213> Artificial Sequence	
	<220> <223> deltaG287NZ-961	
25	<400> 99 Met Ala Ser Pro Asp Val Lys Ser Ala Asp Thr Leu Ser Lys Pro Ala 1 5 10 15	
	Ala Pro Val Val Ser Glu Lys Glu Thr Glu Ala Lys Glu Asp Ala Pro 20 25 30	
30	Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro Ser Ala Gln Gly Gly Gln 35 40 45	
	Asp Met Ala Ala Val Ser Glu Glu Asn Thr Gly Asn Gly Gly Ala Ala 50 55 60	
35	Ala Thr Asp Lys Pro Lys Asn Glu Asp Glu Gly Ala Gln Asn Asp Met 65 70 75 80	
	Pro Gln Asn Ala Ala Asp Thr Asp Ser Leu Thr Pro Asn His Thr Pro 85 90 95	
40	Ala Ser Asn Met Pro Ala Gly Asn Met Glu Asn Gln Ala Pro Asp Ala 100 105 110	
	Gly Glu Ser Glu Gln Pro Ala Asn Gln Pro Asp Met Ala Asn Thr Ala 115 120 125	
45	Asp Gly Met Gln Gly Asp Asp Pro Ser Ala Gly Gly Glu Asn Ala Gly 130 135 140	
	Asn Thr Ala Ala Gln Gly Thr Asn Gln Ala Glu Asn Asn Gln Thr Ala 145 150 155 160	
50	Gly Ser Gln Asn Pro Ala Ser Ser Thr Asn Pro Ser Ala Thr Asn Ser 165 170 175	
	Gly Gly Asp Phe Gly Arg Thr Asn Val Gly Asn Ser Val Val Ile Asp 180 185 190	
55	Gly Pro Ser Gln Asn Ile Thr Leu Thr His Cys Lys Gly Asp Ser Cys 195 200 205	
	Ser Gly Asn Asn Phe Leu Asp Glu Glu Val Gln Leu Lys Ser Glu Phe	

## EP 1 790 660 A2

	210	215	220
5	Glu Lys Leu Ser Asp Ala Asp Lys Ile Ser Asn Tyr Lys Lys Asp Gly 225 230 235 240		
	Lys Asn Asp Gly Lys Asn Asp Lys Phe Val Gly Leu Val Ala Asp Ser 245 250 255		
10	Val Gln Met Lys Gly Ile Asn Gln Tyr Ile Ile Phe Tyr Lys Pro Lys 260 265 270		
	Pro Thr Ser Phe Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg Arg Ser 275 280 285		
15	Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp Thr Leu 290 295 300		
	Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly Asn Ile 305 310 315 320		
	Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala Glu Lys 325 330 335		
20	Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro Ser Lys 340 345 350		
	Gly Glu Met Leu Ala Gly Thr Ala Val Tyr Asn Gly Glu Val Leu His 355 360 365		
25	Phe His Thr Glu Asn Gly Arg Pro Ser Pro Ser Arg Gly Arg Phe Ala 370 375 380		
	Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile Asp Ser 385 390 395 400		
30	Gly Asp Gly Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala Ile Asp 405 410 415		
	Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Gly Asp Val 420 425 430		
35	Ser Gly Lys Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly Lys Tyr 435 440 445		
	Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Phe Gly Val Phe Ala 450 455 460		
40	Gly Lys Lys Glu Gln Asp Gly Ser Gly Gly Ala Thr Asn Asp 465 470 475 480		
	Asp Asp Val Lys Ala Ala Thr Val Ala Ile Ala Ala Ala Tyr Asn 485 490 495		
45	Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp 500 505 510		
	Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp Ala Thr Ala Ala Asp 515 520 525		
50	Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu Lys Lys Val Val Thr 530 535 540		
	Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys 545 550 555 560		
55	Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala 565 570 575		
	Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala Ala Leu Asp Ala Thr - -		

## EP 1 790 660 A2

	580	585	590	
5	Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu 595 600 605			
	Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val 610 615 620			
10	Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp 625 630 635 640			
	Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala 645 650 655			
	Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala 660 665 670			
15	Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Ala 675 680 685			
	Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys 690 695 700			
20	Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn Lys Asp Asn Ile Ala 705 710 715 720			
	Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg Glu Glu Ser Asp Ser 725 730 735			
25	Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr Thr Glu Lys Leu Asp 740 745 750			
	Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala Asp His Asp Thr Arg 755 760 765			
30	Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu Arg Lys Glu Thr Arg 770 775 780			
	Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu Phe Gln Pro Tyr 785 790 795 800			
35	Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val Gly Gly Tyr Lys Ser 805 810 815			
	Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg Phe Thr Glu Asn Phe 820 825 830			
40	Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser Ser Gly Ser Ser Ala 835 840 845			
	Ala Tyr His Val Gly Val Asn Tyr Glu Trp 850 855			
45	<210> 100 <211> 4425 <212> DNA <213> Artificial Sequence			
	<220> <223> deltaG983-ORF46.1			
50	<400> 100 atgacttctg cggccgactt caatgcaggc ggtaccggta tcggcagcaa cagcagagca acaacagcga aatcagcgc agtatcttc gccggtatca agaacgaaat gtgcaaagac agaagcatgc tctgtgccgg tcgggatgac gttgcggta cagacagggc tgccaaatgc aatgcggccccc cccccgaatct gcataccggc gactttccaa acccaaatgc cgcataacaag aatttgatca acctcaaacc tgcaattgaa gcaggctata caggacgcgg ggttagaggt ggtatcgtcg acacaggcga atccgtcggc agcatatcct ttcccgaaact gtatggcaga aaagaacacg gctataacga aaattacaaa aactatacgg cgtatatgcg gaaggaagcg	60 120 180 240 300 360 420		

	cctgaagacg gaggcggtaa agacattgaa gcttcttcg acgatgaggc cgttatagag	480
5	actgaagcaa agccgacgga tatccgcac gtaaaaagaaa tcggacacat cgattggtc	540
	tcccataatta ttggcgccg ttccgtggac ggcagacctg caggcggtat tgcgccgat	600
	gcgacgctac acataatgaa tacgaatgat gaaaccaaga acgaaatgat gttgcagcc	660
	atccgcatacg catggctaa gctggcggaa cgtggcgtgc gcatcgtaa taacagttt	720
	ggaacaacat cgaggcggc cactggcgc ctttccaaa tagccaattt ggaggagcag	780
	taccgccaag cgttgcgtca ctattccgcg ggtataaaaa cagacgaggg tatccgcctg	840
	atgcaacaga ggcattacgg caacctgtcc taccacatcc gtaataaaaa catgccttc	900
10	atctttcga caggcaatga cgcacaagct cagccaaaca catatgccct attgcattt	960
	tatgaaaaag acgctcaaaa aggcatatc acagtcgcag gcgttagaccg cagtggagaa	1020
	aagttcaaac gggaaatgta tggagaacccg ggtacagaac cgcttgagta tggcttcaac	1080
	cattgcggaa ttatgcctat gtggtcctg tcggcacccct atgaacgaaag cgtccgttcc	1140
	acccgtacaa acccgattca aattccgcga ataccctttt ccgcacccat cgtaaccggc	1200
	acggcggctc tgctgtgc gaaataccgg tggatgagca acgacaaccc gcttaccacg	1260
	ttgctgacga cggctcagga catcggtgca gtcggcgtgg acagcaagtt cggctgggaa	1320
15	ctgctggatg cgggttaaggc catgaacgga cccgcgtccct ttccgttccg cgacttacc	1380
	gccgatacga aaggtaatccg cgtatattccg tactccctcc gtaacgcacat ttcaaggcag	1440
	ggcggcctga tcaaaaaagg cggcggccaa ctgcactgc acggcaacaa cacctatacg	1500
	ggcaaaaacca ttatcgaagg cgttgcgtc gtgttgcgt gcaacaacaa atcgatatg	1560
	cgcgtcgaaa ccaaaagggtgc gctgatttat aacggggcgg catccggcgg cagcctgaac	1620
	agcgacggca ttgtctatct ggcagatacc gaccatccg ggcggaaacgaa aaccgtacac	1680
20	atcaaaggca gtctgcagct ggacggcaaa ggtacgcgt acacacgttt gggcaactg	1740
	ctgaaagtgg acggtagccgc gattatcgcc ggcggactgt acatgtcgcc acgcggcaag	1800
	ggggcaggct atctcaacag taccggacga cgttccctt tcctgtgtc cgccaaaatc	1860
	ggcaggatcttccca cacaacatc aaacccgacg gggccctgtc ggcttccctc	1920
	gacagcgtcg aaaaaacacg gggcagtgaa ggcacacgc tgcgttccatc tgccgtcgc	1980
	ggcaatgcgg cacgactgc ttcggcagcg gcacattccg ccgcggccgg tctgaaacac	2040
	gccgtagaac agggcggcag caatctggaa aacctgtatgg tgcactgtgg tgcctccgaa	2100
25	tcatccgaa caccggagac ggttggaaact gcggcagccg accgcacaga tatgcgggc	2160
	atccgccttcc acggcgcac ttccgcgc ggcggcagccg tacagcatgc gaatgcggcc	2220
	gacgggtgtac gcataatccaa cagtctcgcc gctaccgtct atgcgcacag taccggccgc	2280
	catgcccata tgcaggggacg ccgcctgaaac gccgtatcg acgggttggg ccacaacggc	2340
	acgggtctgc ggcgtatcgc gcaaacccaa caggacggtg gaaacgtggg acagggcgt	2400
	gttgaaggca aaatgcgcgg cagtacccaa accgtcggca ttgcccgcgaa aaccggcggaa	2460
30	aatacgcacg cagccgcac actggcgtatc ggacgcacgca catggagcga aaacagtgc	2520
	aatgcaaaaaa ccgcacgcatt tagtctgttt gcaggcatac gcgcacgtgc gggcgtatcc	2580
	ggctatcttc aaggctgtt ctccatcgca cgtctccaaa acagcatcgc ccgcacgc	2640
	ggtgcggacg aacatgcggg aggcaacgc taaacggcactc tgatgcgtc gggcgcactg	2700
	ggcgggtgtca acgttccgtt tgccgcacgc ggagatttga cggtcgaaagg cggctcgcgc	2760
	tacgaccctgc tcaaacagga tgcattcgcc gaaaaaggca gtgtttttggg ctggagcggc	2820
	aacagccta ctgaaaggcac gtcggcggcga ctcgcgggtc tgaagctgtc gcaacccttgc	2880
35	agcgataaaag cgttccgtt tgcacacggcg ggcgttggaaac ggcacccatc cggacgcgac	2940
	tacacggtaa cggggcggctt taccggcgcg actgcacca cccggcggac gggggcacgc	3000
	aatatgcgcg acaccgcgtt ggttgcggcgc ctggggcggg atgtcgaatt cggcaacggc	3060
	tggAACGGCT tggcacgtt cagctacgcc ggttccaaac agtacggcaa ccacagcggaa	3120
	cgagtgcggcg taggttaccgc gttccctcgac ggtggcggag gcactggatc ctcagatttg	3180
	gcaaaacgatt cttttatccg gcagggttctc gaccgtcgcg atttcgaacc cgacgggaaa	3240
	taccacccat tcggcggcgg gggggaaactt ggcggcgc gggccatcat cggattggg	3300
40	aaaatacacaac gccatcgtt gggcaacccgtt atgatcaac aggccggccat taaaggaaat	3360
	atcggttaca ttgtccgtt ttccgcacatc gggcacgcac tccattcccc cttgcacaac	3420
	catgcctcac attccgattt tgatgcgtt ggtatcccc tgcacggatt tagcctttac	3480
	cgcacccattt gggacggata cgaacaccat cccggcgcacg gctatgcacgg gccacagggc	3540
	ggcggctata ccgcgtccaa aggccgcggg gatataataca gctacgcacat aaaaggcggt	3600
45	gccccaaaat tccgcctcaa cctgtacccgcg aaccgcacgc cccggacacgc gttggccgcac	3660
	cgtttccaca atgcggtagt tatgtcgcg acaggatgcg ggcacgcgtt caaacgcgc	3720
	acccgatatac gccccggagct ggacagatgcg ggcaatgcgc cccggacacgc caacggcact	3780
	gcagatatacg taaaacatc catggcgcgc gcaaggagaaa ttgtcggcgc aggcgtatgc	3840
	gtgcaggggca taagcgttgcg cttcaacatt gctgtcatgc acggcttggg tctgcttcc	3900
	accggaaaaca agatggcgcg catcaacgtt ttggcgcacact taaaggactat	3960
	gccgcggcagc ccatccgcga ttggcgcgtt caaaacccca atgcgcacaca aggcatagaa	4020
50	gcgcgtcgcg atatctttat ggcgcgttcc cccatccaaat ggttggagc ttttgcgggaa	4080
	aaatacgcgt tggcgccat cccggcgcacat cctatcaacgc ggtgcgcacat gggcgcgtatc	4140
	gcattggcga aaggaaatc cccggcgtcgc gacaatttttgc cccatccaaat ggttggagc	4200
	tacccgtccc cttaaccattt cccggcgttcc cgttccaaactt tggagcgc gttacggcaaa	4260
	gaaaacatca cctcctcaac cgtggcgcgc tcaaaacggca aaaaatgtcaa actggcagac	4320
	caacggccacc cgaagacagg cgtaccgtt gacggtaaag ggttccgaa ttttggaaag	4380
55	cacgtgaaat atgatcgtt cccggcaccac caccaccacc actga	4425

<211> 1474  
 <212> PRT  
 <213> Artificial Sequence  
 5  
 <220>  
 <223> deltaG983-ORF46.1  
 <400> 101  
 Met Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser  
 1 5 10 15  
 10 Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly  
 20 25 30  
 Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg  
 35 40 45  
 15 Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro  
 50 55 60  
 Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys  
 65 70 75 80  
 20 Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg  
 85 90 95  
 Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile  
 100 105 110  
 25 Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn  
 115 120 125  
 Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly  
 130 135 140  
 30 Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu  
 145 150 155 160  
 Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His  
 165 170 175  
 35 Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg  
 180 185 190  
 Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr  
 195 200 205  
 40 Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala  
 210 215 220  
 Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe  
 225 230 235 240  
 45 Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn  
 245 250 255  
 Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp  
 260 265 270  
 50 Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn  
 275 280 285  
 Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr  
 290 295 300  
 55 Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe  
 305 310 315 320  
 Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp

	325	330	335
5	Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr 340 345 350		
	Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp 355 360 365		
10	Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn 370 375 380		
	Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly 385 390 395 400		
15	Thr Ala Ala Leu Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn Asp Asn 405 410 415		
	Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly 420 425 430		
20	Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys Ala Met 435 440 445		
	Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp Thr Lys 450 455 460		
25	Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser Gly Thr 465 470 475 480		
	Gly Gly Leu Ile Lys Lys Gly Ser Gln Leu Gln Leu His Gly Asn 485 490 495		
30	Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Ser Leu Val Leu 500 505 510		
	Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly Ala Leu 515 520 525		
35	Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp Gly Ile 530 535 540		
	Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His 545 550 555 560		
40	Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg 565 570 575		
	Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys 580 585 590		
45	Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr 595 600 605		
	Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr 610 615 620		
50	Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala Ser Leu 625 630 635 640		
	Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu Ser Tyr 645 650 655		
55	Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala His 660 665 670		
	Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln Gly Gly Ser Asn 675 680 685		
	Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu Ser Ser Ala Thr		

## EP 1 790 660 A2

	690	695	700
5	Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg	705                    710	Thr Asp Met Pro Gly
			715                    720
	Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Val Gln His	725                    730	735
10	Ala Asn Ala Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala Ala Thr	740                    745	750
	val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly Arg Arg	755                    760	765
	Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly Leu Arg	770                    775	780
15	Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln Gly Gly	785                    790	800
	Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala	805                    810	815
20	Lys Thr Gly Glu Asn Thr Thr Ala Ala Ala Thr Leu Gly Met Gly Arg	820                    825	830
	Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser	835                    840	845
25	Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys	850                    855	860
	Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr	865                    870	880
30	Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln	885                    890	895
	Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp	900                    905	910
35	Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys Gln Asp Ala	915                    920	925
	Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser Leu Thr	930                    935	940
40	Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln Pro Leu	945                    950	960
	Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val Glu Arg Asp Leu	965                    970	975
45	Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly Ala Thr Ala	980                    985	990
	Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His Thr Arg Leu Val	995                    1000	1005
50	Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly Trp Asn Gly Leu	1010                    1015	1020
	Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn His Ser Gly	1025                    1030	1040
55	Arg Val Gly Val Gly Tyr Arg Phe Leu Asp Gly Gly Gly Thr Gly	1045                    1050	1055
	Ser Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg		

	1060	1065	1070
5	Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg Gly 1075 1080 1085		
	Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln Ser 1090 1095 1100		
10	His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly Asn 1105 1110 1115 1120		
	Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu Val His Ser 1125 1130 1135		
15	Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly Ser 1140 1145 1150		
	Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu 1155 1160 1165		
	His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Tyr Pro 1170 1175 1180		
20	Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val 1185 1190 1195 1200		
	Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln 1205 1210 1215		
25	Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly 1220 1225 1230		
	Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp 1235 1240 1245		
30	Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val 1250 1255 1260		
	Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp Ala 1265 1270 1275 1280		
35	Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu 1285 1290 1295		
	Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala 1300 1305 1310		
40	Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ile Arg Asp Trp 1315 1320 1325		
	Ala Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn 1330 1335 1340		
45	Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly 1345 1350 1355 1360		
	Lys Tyr Gly Leu Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln 1365 1370 1375		
50	Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn 1380 1385 1390		
	Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg 1395 1400 1405		
55	Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr 1410 1415 1420		
	Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp		

1425            1430            1435            1440

5            Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro  
 1445            1450            1455

Asn Phe Glu Lys His Val Lys Tyr Asp Thr Leu Glu His His His His  
 1460            1465            1470

His His

10

<210> 102

<211> 3939

<212> DNA

<213> Artificial Sequence

15

<220>

<223> deltag983-741

<400> 102

atgacttctg cggccgactt caatgcaggc ggtaccggta tcggcagcaa cagcagagca 60

acaacagcga aatcagcagc agtatcttac gcccgttatca agaacaaat gtccaaagac 120

agaagcatgc tctgtgccgg tcgggatgac gttcggttta cagacaggga tgccaaaatc 180

aatggccccc ccccgaaatct gcataccggg gactttccaa accccaaatga cgcatacaag 240

aatttgatcca accttcaaacc tgcaatttggaa gcaggctata caggacgcgg gtagaggta 300

ggtatcgctg acacaggcga atccgtcggc agcatatctt ttcccgaact gtatggcaga 360

aaagaacacg gctataacga aaatttacaaa aactatacgg cgtatatgcg gaaggaagcg 420

cctgaagacg gaggcggtaa agacatttggaa gcttctttcg acgatgaggc cgttataagag 480

actgaagcaa agccgacggg tatccggcac gtaaaaagaaa tcggacacat cgatgggtc 540

tcccatatta ttggcggggc ttccgtggac ggcagacctg caggcggtat tgcccccgtat 600

ggcacgtac acataatgaa tacgaatgtat gaaaccaaga acgaaatgtat ggttcagcc 660

atccgcattt catgggtcaa gctggcgaa ctggcgctgc gcatcgtaa taacagtttt 720

ggaacaacat cgaggcgagg cactgcccac cttttccaaa tagccaattt ggaggagcag 780

taccgccaag ctttgcgtcga ctattccggc ggtgataaaaa cagacgaggg tatccgcctg 840

atgcaacaga gcgattacgg caacctgtcc taccacatcc gtaataaaaa catgcatttc 900

atcttttcg caggcaatccg cgcacaactg cagccccaca catatgcctt attggcattt 960

tatgaaaaaa acgctcaaaa aggcatattt acagtcgcag gctgttagaccg cagtgagaa 1020

aagttcaaac gggaaatgtt tggagaaccg ggtacagaac cgtttagata tggttccaa 1080

cattgcccggaa ttactgcccattt gtttgtccgt tcggcacccct atgaagcaag cgtccgtttc 1140

acccgtacaa acccgattca aattgcccggaa acatcctttt ccgcaccat cgtAACCGGC 1200

acggcggttc tgcgtctgca gaaatcccg tggatgagca acgacaacct gcgtaccacg 1260

ttgctgacg cggctcaggaa catcggtcga gtcggcggtgg acagcaagtt cggctgggaa 1320

ctgctggatg cgggttggggc catgaacggg cccggcgctt ttccgttccgg cactttacc 1380

gccgatcga aaggatcattt cgatatttgcg tactccttcc gtaacgacat ttccaggcag 1440

ggcggcctga tcaaaaaagg cggcaggccaa ctgcaactgc acggcaacaa caccatacg 1500

ggcaaaacca ttatcgaaagg cgggttcgtcgt gtgttgcgtc gcaacaacaa atcgatatg 1560

cgcgctgaaa ccaaagggtgc gctgattttt aacggggcggg catccggcgg cagccgtaaac 1620

agcgacggc ttgtcttatctt ggcagatacc gacaatccg gcgcaaaacga aaccgtacac 1680

atcaaaaggc gtcgtcagttt ggacggccaa ggtacgtgtt acacacgttt gggcaactg 1740

ctgaaagtgg acgggtacggc gattatccggc ggcacactgt acatgtccgc acggcgcaag 1800

ggggcaggct atctcaacag taccggacga cgttcccct tcctgagtgc cgccaaaatc 1860

gggcaggatt attctttctt cacaaacatc gaaaccgacg gcccgtctgt ggcttccctc 1920

gacagcgctg aaaaaaaacgc gggcagtggaa ggcacacgcg tgccttata tgcgtcgcc 1980

ggcaatcgcc caccggactgc ttccggcaggc gcacattccg cggccggccg tctgaaacac 2040

gccgtaaac acggcgccgac caatctggaa aacctgtatgg tcgacttggc tgccctccgaa 2100

tcatccgcaaa caccggagac ggttgaactt gcccggccg accgcacaga tatgcccggc 2160

atccgcctt acggcgcaac ttcccgccca gcccggccg tacagcatgc gaatgcccgc 2220

gacgggtgtac gcatcttcaa cagtctcgcc gctaccgtct atgcccacac taccggcccc 2280

catggccata tgcaggggacg cccgttggaa gcccgtatccg acgggttggc ccacaacggc 2340

acgggtctgc gcttcatttc gcaaaacccaa caggacgggtg gaaacgtggc acagggcggt 2400

gttgaaggca aatgtcgccgg cgttccaaa accgtccggca ttggccggaa aaccggcgaa 2460

aatacgcacg cagccggccac actgggcattt ggacgcggca catggggccaa aaacagtgc 2520

aatgcaaaaa ccgacagcat tagtctgtttt gcaggcatac ggcacgtatc gggcgatatc 2580

ggctatctca aaggcctgtt ctccctacggc cgctacaaaaa acagcatcag cccgaggcacc 2640

ggtgcggacg aacatgcggg aggccggctc aacggcacgc tgatgcactt gggcgactg 2700

ggcgggtgtca acgttccgtt tgccgcacac ggagatttga cggtcgaagg cggctcgcc 2760

tacgaccatgc tcaaacacgga tgcatttgc gaaaaggca gtgttttggg ctggagcgcc 2820

aacagcctca ctgaaaggcactt gctggcggc gtcggccggc tgaagctgtc gcaaccctt 2880

agcgataaaag ccgtccgttt tgcaacggcg ggcgttggaa ggcacctgaa cggacgcgac 2940

tacacggtaa cggcggctt taccggcgactgcagcaa ccggcaagac gggggcacgc 3000  
 aatatgcgc acacccgtct ggttccggc ctggcgcgg atgtcgaatt cgcaacggc 3060  
 5 tggAACGGCT tggCACGTT cagctacGCC ggttccaaac agtacggcaa ccacaggg 3120  
 cgagtgcgc taggctaccg gttccctcgag ggatccggag ggggtgggt cgccgcgc 3180  
 atcggtgtcg ggcttgcga tgcactaacc gcaccgctcg accataaaga caaagg 3240  
 cagtcttga cgctggatca gtccgtcagg aaaaacgaga aactgaagct ggcggcacaa 3300  
 ggtgcggaaa aaacttatgg aaacggtgac agcctaata cgggcaatt gaagaacgac 3360  
 aaggtcagcc gtttcgactt tatccgcca atcgaagtgg acgggcagct cattac 3420  
 gagagtggag agttccaagt atacaacaa agccattccg ccttaaccgc ctttcagacc 3480  
 10 gagcaaatac aagattcggg gcattccggg aagatgggtg cgaaacgcca gttcagaatc 3540  
 ggcgacatag cggcgaaca tacatctt gacaagttc cccaaaggcgg caggcgaca 3600  
 tattcgcgggaa cggcgttccggttcagacat gccggcggaa aactgaccata caccatagat 3660  
 ttccgcggca agcaggaaa cggcaaaatc gaacatttga aatcgccaga actcaatgtc 3720  
 gacctggccg ccggcgatata caagccgat ggaaaacgccc atgccgtcat cagcggttcc 3780  
 gtccttaca accaagccga gaaaggcagt tactccctcg gtatcttgg cgaaaagcc 3840  
 caggaagttg ccggcagcgc ggaagtggaa accgtaaacg gcatacggcca tatcggcctt 3900  
 15 gccgccaagc aactcgagca ccaccaccac caccactga 3939

<210> 103  
 <211> 1312  
 <212> PRT  
 <213> Artificial Sequence

20 <220>  
 <223> deltaG983-741

<400> 103  
 Met Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser  
 1 5 10 15

25 Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly  
 20 25 30

Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg  
 35 40 45

30 Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro  
 50 55 60

Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys  
 65 70 75 80

35 Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg  
 85 90 95

Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile  
 100 105 110

40 Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn  
 115 120 125

Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly  
 130 135 140

45 Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu  
 145 150 155 160

Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His  
 165 170 175

50 Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg  
 180 185 190

Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr  
 195 200 205

55 Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala  
 210 215 220

Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe  
 225 230 235 240  
 5 Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn  
 245 250 255  
 Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp  
 10 260 265 270 275  
 Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn  
 275 280 285  
 Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr  
 290 295 300  
 15 Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe  
 305 310 315 320  
 Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp  
 325 330 335  
 20 Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr  
 340 345 350  
 Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp  
 355 360 365  
 25 Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn  
 370 375 380  
 Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly  
 385 390 395 400  
 30 Thr Ala Ala Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn Asp Asn  
 405 410 415  
 Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly  
 420 425 430  
 35 Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys Ala Met  
 435 440 445  
 Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp Thr Lys  
 450 455 460  
 Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser Gly Thr  
 465 470 475 480  
 40 Gly Gly Leu Ile Lys Lys Gly Ser Gln Leu Gln Leu His Gly Asn  
 485 490 495  
 Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly Ser Leu Val Leu  
 500 505 510  
 45 Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly Ala Leu  
 515 520 525  
 Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp Gly Ile  
 530 535 540  
 50 Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His  
 545 550 555 560  
 Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg  
 565 570 575  
 55 Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys  
 580 585 590

Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr  
 595 600 605  
 5 Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr  
 610 615 620  
 Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala Ser Leu  
 625 630 635 640  
 10 Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu Ser Tyr  
 645 650 655  
 Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala His  
 660 665 670  
 15 Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln Gly Ser Asn  
 675 680 685  
 Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu Ser Ser Ala Thr  
 690 695 700  
 20 Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr Asp Met Pro Gly  
 705 710 715 720  
 Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Val Gln His  
 725 730 735  
 25 Ala Asn Ala Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala Ala Thr  
 740 745 750  
 Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly Arg Arg  
 755 760 765  
 30 Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly Leu Arg  
 770 775 780  
 Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln Gly Gly  
 785 790 795 800  
 35 Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala  
 805 810 815  
 Lys Thr Gly Glu Asn Thr Ala Ala Ala Thr Leu Gly Met Gly Arg  
 820 825 830  
 Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser  
 835 840 845  
 40 Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys  
 850 855 860  
 Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr  
 865 870 875 880  
 45 Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln  
 885 890 895  
 Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp  
 900 905 910  
 50 Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys Gln Asp Ala  
 915 920 925  
 Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser Leu Thr  
 930 935 940  
 55 Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln Pro Leu  
 945 950 955 960

Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val Glu Arg Asp Leu  
 965 970 975  
 5 Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly Ala Thr Ala  
 980 985 990  
 Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His Thr Arg Leu Val  
 995 1000 1005  
 10 Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly Trp Asn Gly Leu  
 1010 1015 1020  
 Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn His Ser Gly  
 1025 1030 1035 1040  
 15 Arg Val Gly Val Gly Tyr Arg Phe Leu Glu Gly Ser Gly Gly Gly  
 1045 1050 1055  
 Val Ala Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala Pro  
 1060 1065 1070  
 20 Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln Ser  
 1075 1080 1085  
 Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln Gly Ala Glu Lys  
 1090 1095 1100  
 25 Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys Leu Lys Asn Asp  
 1105 1110 1115 1120  
 Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu Val Asp Gly Gln  
 1125 1130 1135  
 30 Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr Lys Gln Ser His  
 1140 1145 1150  
 Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln Asp Ser Glu His  
 1155 1160 1165  
 Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile Gly Asp Ile Ala  
 1170 1175 1180  
 35 Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala Thr  
 1185 1190 1195 1200  
 Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu Thr  
 1205 1210 1215  
 40 Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Ile Glu His  
 1220 1225 1230  
 Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala Asp Ile Lys  
 1235 1240 1245  
 45 Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser Val Leu Tyr Asn  
 1250 1255 1260  
 Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe Gly Gly Lys Ala  
 1265 1270 1275 1280  
 50 Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val Asn Gly Ile Arg  
 1285 1290 1295  
 His Ile Gly Leu Ala Ala Lys Gln Leu Glu His His His His His His  
 1300 1305 1310  
 55 <210> 104  
 <211> 4344

<212> DNA  
<213> Artificial Sequence

5 <220>  
<223> deltaG983-961

<400> 104

atgacttctg	cggccgactt	caatgcaggc	ggtaccggta	tcggcagcaa	cagcaagac
acaacagcg	aatcagcgc	agtatcttc	gccggtatca	agaacgaaat	gtgcaagac
agaagctgc	tctgtccgg	tcgggatgac	gttgcggta	cagacaggga	tgccaaaatc
aatggcccc	ccccgaatct	gcataccgga	gactttccaa	acccaaatga	cgcataacaag
aatttgatca	acctcaaacc	tgcaatttga	gcaggctata	caggacggg	ggttagggta
ggtatcg	acacaggcga	atccgtcg	agcatatcct	ttcccgaact	gtatggcaga
aaagaacacg	gctataacga	aaattacaaa	aactatacgg	cgatatatgc	gaaggaagcg
cctgaagacg	gaggcggtaa	agacattgaa	gcttcttgc	acgatgaggc	cgttatagag
actgaagcaa	agccgacgga	tatccgcac	gtaaaaagaaa	tcggacacat	cgatttggc
tcccattat	ttggcggcg	ttccgtggac	ggcagacctg	caggcggat	tgcggccgat
gcgacgtac	acataatgaa	tgcaatgtat	gaaaccaaga	acgaaatgat	gttgcaggc
atccgcaatg	catgggtcaa	gctggcgaa	cgtggcgtgc	gcatcgtcaa	taacagttt
ggaacaacat	cgagggcagg	cactgcccac	ctttccaaa	tagccaattc	ggaggagcag
taccgccaag	cgttgc	ctattccggc	ggtgataaaa	cagacgaggg	tatccgcctg
atgcaacaga	gcfattacgg	caacctgtcc	taccacatcc	gtaataaaaa	catgcgtttc
atcttttgc	caggcaatga	cgcacaagct	cagcccaaca	catatgcct	attgcattt
atatggaaaag	acgctaaaaa	aggcattac	acagtgcag	gctgtagacc	cagtggagaa
aagttcaaac	gggaaatgt	tggagaaccc	ggtacagaac	cgctttagta	tggtcttcaac
cattgcgaa	ttactgccc	gtgtgtcc	tcggcaccc	atgaagcaag	cgtcgtttc
acccgtacaa	acccgattca	aatttgcgga	acatccccc	ccgcacccat	cgtaaccggc
acggcggctc	tgctgtgc	gaaataccc	tggatgagca	acgacaacct	gcgtaccacg
ttgtctgacg	cgggtcagga	catcgtgtca	gtcgcgtgg	acagcaagt	cggctggga
ctgtgttag	cgggtaaaggc	catgaacgg	cccgctcc	ttccgttccg	cgacttacc
gcccatacg	aggatcatc	cgatattgc	tactccccc	gtaacgacat	ttcaggcagc
ggcggctg	tcaaaaaagg	cggcagccaa	ctgcaactgc	acggcaacaa	cacctatacg
ggcaaaacca	ttatcgagg	cggttcgtg	gtgttgtacg	gcaacaacaa	atcggtatag
cgcgtcgaa	ccaaagggtgc	gctgatttat	aacggggcgg	catccggcgg	cagccgtaaac
agcgtacgca	ttgtctatct	ggcagatacc	gaccatcc	gkgcaacacg	accgtacac
atcaaaaggc	gtctcgatc	ggacggcaaa	ggtaacgt	acacacgtt	ggcgaactg
ctgaaagtgg	acgttacggc	gattatcg	ggcaactgt	atcgatcg	acgcggcag
ggggcaggct	atctcaacag	taccggacca	cgtgttcc	tcctgagtgc	cgccaaaatc
gggcaggatt	attctttctt	cacaaacatc	gaaaccgac	gcccctgt	ggcttccctc
gacagcg	aaaaaacagc	gggcagtgaa	ggcgcacac	tgtcttatt	tgtccgtcgc
ggcaatcg	cacggactgc	ttcggcagcg	gcacattcc	cgcccggc	tctgaaacac
gccctagaac	aggggcggcag	caatctggaa	aacctgtat	tcgaacttgc	tgctcccgaa
tcatccgca	cacccggagac	ggttttaaact	gcccgcgg	accgcacaga	tatgcggcgc
atccgcccc	acggcgcaac	tttccgcgca	gcccgcgg	tacagcatgc	aatggccgccc
gacggtgtac	gcatttca	cagtctcg	gctaccgt	atgcccacag	taccggcgcc
catggccata	tgcagggac	ccgcctgaaa	gccgtatcg	acgggttgg	ccacaacggc
acgggtctgc	gcgtcatgc	gcaaaacccaa	caggacgg	gaacgttgg	acaggcgggt
gttgaaggca	aaatgcgcgg	cagtacccaa	accgtggc	ttgcgcgaa	aaccggcggaa
aatacgcac	caggcccac	actgggcat	ggacgcac	atggagcga	aaacagtgc
aatgcaaaaa	ccgacagcat	tagtctgtt	gcaggcatac	ggcacatgc	ggcgcataatc
ggctatctc	aaggcctgtt	ctcttacgg	cgctacaaa	acagcatcg	ccgcagcacc
ggtgccgac	aacatgcgga	aggcagcgtc	aacggcac	tgatgcacgt	gggcgcactg
ggccgtgtca	acgttccgtt	tgccgc	ggagatttga	cggtcgaagg	cggtctgcgc
tacgacc	tcaaaacagg	tgatttcg	gaaaaaggca	gtgttttgg	ttggagcggc
cgagtcgtc	ctgaaggc	gctgggtcg	ctcgccgg	tgaagctgtc	gcaacccctt
agcgataaa	ccgttctgtt	tgcaacggc	ggcgttgg	ggcacctgaa	cggacgcgcac
tacacggtaa	cgggcggctt	taccggc	actgcag	ccggcaagac	gggggcacgc
aatatgcgc	acaccctgtt	ggttgcggc	ctgggc	atgtcgaatt	cggcacacggc
tggaaacgct	tggcacgtt	cagctac	ggtttccaaac	agtacggca	ccacacggga
cgagtcgtc	taggttac	gttccctcg	ggttgcgg	gttgcgttgg	gttgcgttgg
gacgacgt	ttaaaaagc	tgcctactgt	gcccatt	gactggat	cgccataaaac
gaaatcaac	gtttcaaaagc	tggagagac	gttgcgtt	ctgcctacaa	caatggccaa
acccaaaaaa	acgcacactgc	agccgtat	atctacgaca	tgtatgaaga	cggcacaaatt
aaaaaaagt	tgactaacct	gaccaaaacc	gaagccgac	actttaaagg	tctgggtctg
aaagtaaaag	ctgcagaatc	tgaaataga	gtcaatgaa	acaaacaaaa	cgtcgatgccc
ggccgtttag	cagatactga	tgccgctctg	aaagttaacaa	ccaagttagc	agacactgtat
ggagaaaaata	taacgcac	tgctgttag	gatgcacca	ccaacgcctc	gaataaatttgc
aatttagaag	cgttggtctg	taccgtcg	actaagacaa	atatcgtaaa	aattgtatgaa
gattcattqq	atgaaaccaa	cactaaggc	aagcatggc	aagcattca	cgatcatgc
			gacgaagcc	tcaaaaccgc	caatgaagcc

aaacagacgg ccgaagaaaac caaacaaaac gtcgatgcca aagtaaaagc tgcagaaact 3780  
 gcagcaggca aagccgaagc tgccgctggc acagctaata ctgcagccga caaggccgaa 3840  
 gctgtcgctg caaaagttaac cgacatcaa gctgatatcg ctacgaacaa agataatatt 3900  
 gctaaaaaaag caaacagtgc cgacgtgtac accagagaag agtctgacag caaatttgtc 3960  
 agaatttgatg gtctgaacgc tactaccgaa aaattggaca cacgcttggc ttctgctgaa 4020  
 aaatccattg ccgatcacga tactcgctg aacgggttgg ataaaaacagt gttagacactg 4080  
 cgaaagaaa cccgccaagg ccttgcagaa caagccgc tctccgtct gttccaacct 4140  
 tacaacgtgg gtccggttcaa tgtaacggct gcagtccggcg gctacaataa cgaatcggca 4200  
 gtcgccatcg gtaccggctt ccgcttacc gaaaacttt ccgccaaagc aggcgtggca 4260  
 gtcggcactt cgtccggttc ttccgcagcc taccatgtcg gctcaatta cgagtggctc 4320  
 gagcaccacc accaccacca ctga 4344

5 <210> 105  
 <211> 1447  
 <212> PRT  
 10 <213> Artificial Sequence

15 <220>  
 <223> deltaG983-961

20 <400> 105  
 Met Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser  
 1 5 10 15

Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly  
 20 25 30

25 Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg  
 35 40 45

Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro  
 50 55 60

30 Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys  
 65 70 75 80

Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg  
 85 90 95

35 Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile  
 100 105 110

Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn  
 115 120 125

40 Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly  
 130 135 140

Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu  
 145 150 155 160

Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His  
 165 170 175

45 Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg  
 180 185 190

Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr  
 195 200 205

50 Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala  
 210 215 220

Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe  
 225 230 235 240

55 Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn  
 245 250 255

Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp  
 260 265 270  
 5 Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn  
 275 280 285  
 Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr  
 290 295 300  
 10 Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe  
 305 310 315 320  
 Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp  
 325 330 335  
 15 Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr  
 340 345 350  
 Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp  
 355 360 365  
 20 Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn  
 370 375 380  
 Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly  
 385 390 395 400  
 25 Thr Ala Ala Leu Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn Asp Asn  
 405 410 415  
 Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly  
 420 425 430  
 30 Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys Ala Met  
 435 440 445  
 Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp Thr Lys  
 450 455 460  
 35 Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser Gly Thr  
 465 470 475 480  
 Gly Gly Leu Ile Lys Lys Gly Gly Ser Gln Leu Gln Leu His Gly Asn  
 485 490 495  
 40 Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly Ser Leu Val Leu  
 500 505 510  
 Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly Ala Leu  
 515 520 525  
 Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp Gly Ile  
 530 535 540  
 45 Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His  
 545 550 555 560  
 Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg  
 565 570 575  
 50 Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys  
 580 585 590  
 Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr  
 595 600 605  
 55 Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr  
 610 615 620

Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala Ser Leu  
 625 630 635 640  
 5 Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu Ser Tyr  
 645 650 655  
 Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala His  
 660 665 670  
 10 Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln Gly Gly Ser Asn  
 675 680 685  
 Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu Ser Ser Ala Thr  
 690 695 700  
 15 Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr Asp Met Pro Gly  
 705 710 715 720  
 Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Ala Val Gln His  
 725 730 735  
 20 Ala Asn Ala Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala Ala Thr  
 740 745 750  
 Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly Arg Arg  
 755 760 765  
 25 Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly Leu Arg  
 770 775 780  
 Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln Gly Gly  
 785 790 795 800  
 30 Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala  
 805 810 815  
 Lys Thr Gly Glu Asn Thr Ala Ala Ala Thr Leu Gly Met Gly Arg  
 820 825 830  
 35 Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser  
 835 840 845  
 Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys  
 850 855 860  
 40 Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr  
 865 870 875 880  
 Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln  
 885 890 895  
 45 Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp  
 900 905 910  
 Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys Gln Asp Ala  
 915 920 925  
 50 Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser Leu Thr  
 930 935 940  
 Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln Pro Leu  
 945 950 955 960  
 55 Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val Glu Arg Asp Leu  
 965 970 975  
 Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly Ala Thr Ala  
 980 985 990

Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His Thr Arg Leu Val  
 995 1000 1005  
 5 Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly Trp Asn Gly Leu  
 1010 1015 1020  
 Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn His Ser Gly  
 1025 1030 1035 1040  
 10 Arg Val Gly Val Gly Tyr Arg Phe Leu Glu Gly Gly Gly Thr Gly  
 1045 1050 1055  
 Ser Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile  
 1060 1065 1070  
 15 Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly  
 1075 1080 1085  
 Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp  
 1090 1095 1100  
 20 Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu  
 1105 1110 1115 1120  
 Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln  
 1125 1130 1135  
 25 Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu  
 1140 1145 1150  
 Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala  
 1155 1160 1165  
 30 Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile  
 1170 1175 1180  
 Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu  
 1185 1190 1195 1200  
 35 Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe  
 1205 1210 1215  
 Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu  
 1220 1225 1230  
 40 Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys  
 1235 1240 1245  
 Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys  
 1250 1255 1260  
 45 Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu  
 1265 1270 1275 1280  
 Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn  
 1285 1290 1295  
 50 Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg  
 1300 1305 1310  
 Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr  
 1315 1320 1325  
 Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala  
 1330 1335 1340  
 55 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu  
 1345 1350 1355 1360

Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly  
 1365 1370 1375  
 5 Leu Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val  
 1380 1385 1390  
 Gly Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg  
 1395 1400 1405  
 10 Phe Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser  
 1410 1415 1420  
 Ser Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp Leu  
 1425 1430 1435 1440  
 15 Glu His His His His His His  
 1445  
 <210> 106  
 <211> 4179  
 <212> DNA  
 20 <213> Artificial Sequence  
 <220>  
 <223> deltag983-961c  
 <400> 106  
 atgacttctg cggccgactt caatgcaggc ggtaccggta tcggcagcaa cagcagagca  
 25 acaacagcga aatcagcagc agtatcttac gccggatca agaacaaat gtccaaagac  
 agaagcatgc tcgtgccgg tcgggatgc gttcggtta cagacaggga tgccaaaatc  
 aatgcccccc ccccgaaatct gcataccgga gactttccaa acccaaatga cgcatacaag  
 aatttgcata acctcaaaacc tgcaattgaa gcaggctata caggacgcgg ggttagaggt  
 ggtatcgtcg acacaggcga atccgtcgcc agcatatcct ttcccgaact gtatggaga  
 aaagaacacg gctataacga aaattacaaa aactatacgg cgtatatgcg gaaggaagcg  
 cctgaagacg gaggcggtaa agacattgaa gcttcttcg acgtataggc cggtatagag  
 30 actgaagcaa agccgacggc tatccggcac gtaaaagaaa tcggacacat cgatggtc  
 tcccatatta ttggcgccg ttccgtggc ggcagacctg caggcggtat tgcggccgat  
 ggcacgctac acataatgaa tacgaatgt gaaacaaga acgaaatgtat ggttgagcc  
 atccgcaatg catgggtcaa gctggcgaa cgtggcgtgc gcatcgtaa taacagttt  
 ggaacaacat cgagggcagg cactgcccac ctttccaaa tagccaattc ggaggagcag  
 taccgccaag cgttgcgtca ctattccggc ggtgataaaaa cagacgggg tatccgcctg  
 35 atgcaacaga gcgattacgg caacctgtcc taccacatcc gtaataaaaa catgttttc  
 atctttcga caggcaatgc cgcacaact cagccaaaca catatccctt attgcattt  
 tataaaaaaac acgctaaaaa aggcatatc acagtgcgag gcgttagaccg cagttggagaa  
 aagttcaaac gggaaatgtt tggagaaccg ggtacagaac cgctttagta tggctcaac  
 cattgcggaa ttactgcccgtt gttgtgcctg tcggcaccct atgaagcaag cgtccgttcc  
 40 acccgtaaca acccgattca aattgcggc acatccttt ccgcaccat cgtAACGGC  
 acggcgctc tgctgcgtca gaaataccgg tggatgagca acgacaacct gcgtaccacg  
 ttgctgcga cggctcaggc catcggtca gtcggcggtgg acagcaagtt cggtggggaa  
 ctgctggatg cgggttaaggc catgaacggc cccgcgtctt ttccgttcgg cactttacc  
 gccgatacga aaggtacatc cgatattgcc tactccttcc gtaacgcacat ttcaaggc  
 45 ggcggcctga tcaaaaaagg cggcaggccaa ctgcaactgc acggcaacaa caccatatacg  
 ggcääacca ttatcgaaagg cggttcgtctg gtgttgtac gcaacaacaa atcgatatacg  
 cgcgtcgaaa ccaaagggtgc gctgatttat aacggggcg gatccgtgaac  
 agcgcacggca ttgttatctt ggcagatacc gaccaatctcg ggcacaaacgaaaccgtac  
 atcaaaggca gtcgtcagct ggacggcaaa ggtacgcgtt acacacgtt gggcaactg  
 50 ctgaaaatgg acggtacggc gattatcgcc ggcaagctgt acatgtcgcc acgcggcaag  
 gggcaggctt atctcaacag taccggacga cgttccctt tcctgagtg cgcacaaatc  
 gggcaggattt atctttctt cacaacatc gaaaccgacg gcccggcg cagccgtaaac  
 gacagcgtcg aaaaaacacg gggcgtgaa ggcacacgc tgcgttattt tgccgtcgc  
 ggcacatcgcc cacggactgc ttccgcgcg gcacattccg cggccggccg tctgaaacac  
 55 gccgtacac agggcgccag caatctggaa aacctgtatgg tcgaactgg tgcctccgaa  
 tcatccgcaa caccggagac gttgaaact gcggcagccg accgcacaga tatgcccggc  
 atccgccccctt acggcgcaac ttccgcgcg gcggcagccg tacagcatgc gaatgccg  
 gacgggtgtac gcatcttcaa cagtctcgcc gtcaccgtt atgcgcacag taccggcc  
 catggccata tgcaggggacg ccgcctgaaa gccgtatcg acgggttgg ccacaacggc  
 acgggtctgc gcgtcatcgca gcaaaccggc caggacgggtg gaaacgtggg acaggcggt  
 gttgaaggca aatgcgcggc cagtacccaa accgtcgccg ttggccgcgaa aaccggcgaa  
 aatacgacag cagccgcac actgggcatg ggacgcagca catggagcga aaacagtgc  
 2520

	aatgcaaaaa ccgacagcat tagtctgttt gcaggcatac ggcacgatgc gggcgatatc	2580
5	ggctatctca aaggcctgtt ctcctacgga cgctacaaaa acagcatcg ccgcagcacc	2640
	ggtgcggacg aacatgcgga aggcacggtc aacggcacgc tgatgcagct gggcgactg	2700
	ggcgggtgtca acgttccgtt tgccgcaacg ggagatttgta cggtcgaagg cggctgccc	2760
	tacgacctgc tcaaacagga tgcattcgcg gaaaaaggca gtgctttggg ctggagcggc	2820
	aacagcctca ctgaaaggcac gctggcgggta ctcgcgggtc tgaagctgtc gcaacccttg	2880
	agcgataaaag ccgtcctgtt tgcaacggcg ggcgtggaaac ggcacctgaa cggacgcgac	2940
10	tacacggtaa cggccggctt taccggcgcg actgcagcaa cggcaagac gggggcacgc	3000
	aatatgccgc acacccgtct ggttgcggc ctgggcgcg atgtcgaatt cgccaacggc	3060
	tggAACGGCT tggcacgtta cagctacgccc ggttccaaac agtacggca ccacagcgg	3120
	cggactggcg taggctaccg ttccctcgag ggtggcggag gcactggatc cgcacaaaac	3180
	gacgacatg tttaaaaaagg tgccactgtg gccattgtg ctgcctacaa caatggccaa	3240
	gaaatcaacgc gttaaaaggcc tggagagacc attcacgaca ttgatgaaga cggcacaatt	3300
	acaaaaaaag acgcaactgc agccgatgtt gaagccgacg actttaaagg tctgggtctg	3360
	aaaaaaagtgc tgactaacct gaccaaaaacc gtcaatgaaa acaaacaaaa cgtcgatgcc	3420
15	aaagtaaaag ctgcagaatc tgaaatagaa aagttaaacaa ccaagttagc agacactgat	3480
	gccgccttag cagatactga tgccgcctcg gatgcaccca ccaacgcctt gaataaatttgc	3540
	ggagaaaaata taacgacatt tgctgaagag actaaagacaa atatcgtaaa aattgatgaa	3600
	aaattaaag ccgtggctga taccgtcgac aacgatgccc aagcattcaa cgatatcgcc	3660
	gattcattog atgaaaccac cactaaggca gacgaagccg tcaaaaacgc caatgaagcc	3720
	aaacagacgg ccgaagaaac caaacaacaa gtcgatgcca aagttaaaggc tgcagaaact	3780
20	gcagcaggca aagccgaagc tgccgcctgc acagctaata ctgcagccg caaggccgaa	3840
	gctgtcgctg caaaaaggta cgcacatcaa gctgatatcg ctacgaacaa agataatatt	3900
	gctaaaaaaag caaacagtgc cgacgtgtac accagagaag agtgcacag caaatttgc	3960
	agaatttgatg gtctgaacgc tactaccgaa aaattggaca cacgcttggc ttctgctgaa	4020
	aaatccattg ccgatcacga tactcgcctg aacggtttg ataaaacagt gtcagacctg	4080
	cgaaagaaaa cccgccaagg cttgcagaa caagccgcgc tctccggctc gttccaacct	4140
	tacaacgtgg gtctcgagca ccaccacacc caccactga	4179
25	<210> 107	
	<211> 1392	
	<212> PRT	
	<213> Artificial Sequence	
30	<220>	
	<223> deltaG983-961c	
	<400> 107	
	Met Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser	
	1 5 10 15	
35	Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly	
	20 25 30	
	Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg	
	35 40 45	
40	Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro	
	50 55 60	
	Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys	
	65 70 75 80	
45	Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg	
	85 90 95	
	Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile	
	100 105 110	
50	Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn	
	115 120 125	
	Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly	
	130 135 140	
55	Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu	
	145 150 155 160	
	Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His	

	165	170	175
5	Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg 180 185 190		
	Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr 195 200 205		
10	Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala 210 215 220		
	Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe 225 230 235 240		
15	Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn 245 250 255		
	Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp 260 265 270		
20	Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn 275 280 285		
	Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr 290 295 300		
25	Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe 305 310 315 320		
	Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp 325 330 335		
	Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr 340 345 350		
30	Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp 355 360 365		
	Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn 370 375 380		
35	Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly 385 390 395 400		
	Thr Ala Ala Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn Asp Asn 405 410 415		
40	Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly 420 425 430		
	Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys Ala Met 435 440 445		
45	Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp Thr Lys 450 455 460		
	Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser Gly Thr 465 470 475 480		
50	Gly Gly Leu Ile Lys Lys Gly Gly Ser Gln Leu Gln Leu His Gly Asn 485 490 495		
	Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly Ser Leu Val Leu 500 505 510		
55	Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly Ala Leu 515 520 525		
	Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp Gly Ile		

## EP 1 790 660 A2

	530	535	540
5	Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His 545 550 555 560		
	Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg 565 570 575		
10	Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys 580 585 590		
	Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr 595 600 605		
15	Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr 610 615 620		
	Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala Ser Leu 625 630 635 640		
20	Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu Ser Tyr 645 650 655		
	Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala Ala His 660 665 670		
	Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln Gly Gly Ser Asn 675 680 685		
25	Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu Ser Ser Ala Thr 690 695 700		
	Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr Asp Met Pro Gly 705 710 715 720		
30	Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Ala Val Gln His 725 730 735		
	Ala Asn Ala Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala Ala Thr 740 745 750		
35	Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly Arg Arg 755 760 765		
	Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly Leu Arg 770 775 780		
40	Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln Gly Gly 785 790 795 800		
	Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala 805 810 815		
45	Lys Thr Gly Glu Asn Thr Thr Ala Ala Ala Thr Leu Gly Met Gly Arg 820 825 830		
	Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser 835 840 845		
50	Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys 850 855 860		
	Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr 865 870 875 880		
55	Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln 885 890 895		
	Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp		

## EP 1 790 660 A2

	900	905	910
5	Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys Gln Asp Ala 915 920 925		
	Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser Leu Thr 930 935 940		
10	Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln Pro Leu 945 950 955 960		
	Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val Glu Arg Asp Leu 965 970 975		
15	Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly Ala Thr Ala 980 985 990		
	Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His Thr Arg Leu Val 995 1000 1005		
20	Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly Trp Asn Gly Leu 1010 1015 1020		
	Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn His Ser Gly 1025 1030 1035 1040		
	Arg Val Gly Val Gly Tyr Arg Phe Leu Glu Gly Gly Gly Thr Gly 1045 1050 1055		
25	Ser Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile 1060 1065 1070		
	Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly 1075 1080 1085		
30	Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp 1090 1095 1100		
	Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu 1105 1110 1115 1120		
35	Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln 1125 1130 1135		
	Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu 1140 1145 1150		
40	Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala 1155 1160 1165		
	Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile 1170 1175 1180		
45	Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu 1185 1190 1195 1200		
	Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe 1205 1210 1215		
50	Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu 1220 1225 1230		
	Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys 1235 1240 1245		
55	Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys 1250 1255 1260		
	Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu		

	1265	1270	1275	1280			
5	Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn 1285 1290 1295						
	Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg 1300 1305 1310						
10	Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr 1315 1320 1325						
	Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala 1330 1335 1340						
15	Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu 1345 1350 1355 1360						
	Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly 1365 1370 1375						
20	Leu Phe Gln Pro Tyr Asn Val Gly Leu Glu His His His His His 1380 1385 1390						
	<210> 108						
	<211> 1947						
	<212> DNA						
25	<213> Artificial Sequence						
	<220>						
	<223> deltaG741-961						
	<400> 108						
30	atggtcgccc	ccgacatcg	tgcggggctt	gccgatgcac	taaccgcacc	gctcgaccat	60
	aaagacaag	gtttcagtc	tttgacgctg	gatcagtccg	tcaaaaaaaa	cgagaaaactg	120
	aaggctggcg	cacaagggtc	ggaaaaaaact	tatggaaacg	gtgacagcct	caatacgggc	180
	aaatttgaaga	acgacaagggt	cagccgttcc	gactttatcc	gccaaatcga	agtggacggg	240
	cagctcatta	ccttggagag	tggagatgc	caagtataca	aacaaagcca	ttccgccta	300
	accgccttcc	agaccggagca	aatacaagat	tcggagcatt	ccggaaagat	gtttgcgaaa	360
	cgccagttca	gaatcggcga	catagcgggc	gaacatacat	cttttgacaa	gtttcccgaa	420
	ggcggcaggg	cgacatatcg	cgggacggcg	ttcggttcag	acgatgcgg	cggaaaactg	480
35	acctacacca	tagatttgc	cgccaaagcag	ggaaacggca	aaatcgaaca	tttggaaatcg	540
	ccagaactca	atgtcgaccc	ggccgcgcg	gatatcaagc	cggatggaaa	acgcccatgcc	600
	gtcatacgcg	gttccgtcct	ttacaacaa	gcccggaaaag	gcagttactc	cctcggtatc	660
	tttggcggaa	aaggcccgag	agttgcggc	agcgcggaaag	tggaaaacgt	aaacggcata	720
	cgcctatcg	gccttgcgc	caagcaactc	gagggtggcg	gaggcactgg	atccgccaca	780
	aacgacgacg	atgttaaaaa	agctgcact	gtggccattt	ctgctgccta	caacaatggc	840
40	caagaaatca	acggtttcaa	agctggagag	accatctacg	acattgtatg	agacggcaca	900
	attacaaa	aagacgcaac	tgcagccat	gttgaagccg	acgactttaa	aggtctgggt	960
	ctgaaaaaa	tctgtactaa	cctgacaaa	accgtcaatg	aaaacaaca	aaacgtcgat	1020
	gc当地aaat	aagctcgaga	atctgaaata	gaaaagttaa	caaccaagtt	agcagacact	1080
	gatgcgcctt	tagcagatac	tgtatgcgc	ctggatgc	ccaccaacgc	cttgaataaa	1140
	ttgggagaaa	atataacgac	atttgctgaa	gagactaaga	caaataatcg	aaaattgtat	1200
45	gaaaaattag	aagccgtggc	tgtatccgt	gacaaggcat	ccgaaggatt	caacgatatc	1260
	gccgattcat	tggatggaaac	caacactaag	gcagacggag	ccgtcaaaac	cgccaaatgaa	1320
	gc当地aaat	cggccgaaga	aaccaaaca	aacgtcgat	ccaaagtaaa	agctgcagaa	1380
	actgcagcg	gc当地aaat	agctgcgc	ggcacagcta	atactgcgc	cgacaaaggcc	1440
	gaagctgtcg	ctgaaaaagt	taccgacatc	aaagctgata	tcgctacaa	caaagataat	1500
	attgtctaaa	aagcaaaacag	tgcgcacgt	tacaccagag	aagagtctga	cagcaaattt	1560
	gtcagaattt	atgttctgaa	cgctactacc	gaaaaattgg	acacacgctt	ggcttctgct	1620
	gaaaaatttca	ttggccgtca	cgatactcgc	ctgaacgggt	tggataaaac	agtgtcagac	1680
50	ctgc当地aaat	aaacccgcca	aggccttgc	gaacaaggcc	cgctctccgg	tctgttccaa	1740
	ccttacaacg	tgggtcggtt	caatgttaacg	gctgcagtc	gcggctacaa	atccgaatcg	1800
	gcagtcgcca	tcggtaccgg	cttccgtt	accgaaaact	ttgcccacaa	agcaggcgtg	1860
	gcagtcgcca	cttcgtccgg	ttcttccgca	gccttaccatg	tcggcgtcaa	ttacgagtg	1920
	ctcgagcacc	accaccacca	ccactg				1947
55	<210>	109					
	<211>	648					

<212> PRT  
<213> Artificial Sequence

<220>  
<223> deltaG741-961

<400> 109  
Met Val Ala Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala  
1 5 10 15

10 Pro Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln  
20 25 30

Ser Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln Gly Ala Glu  
35 40 45

15 Lys Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys Leu Lys Asn  
50 55 60

Asp Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu Val Asp Gly  
65 70 75 80

20 Gln Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr Lys Gln Ser  
85 90 95

His Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln Asp Ser Glu  
100 105 110

25            His Ser Gly Lys Met Val Ala Lys Arg Gin Phe Arg Ile Gly Asp Ile  
              115                            120    125

Ala Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala  
130 135 140

30 Thr Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu  
145 150 155 160

Thr Tyr Thr Thr Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Thr Glu  
165 170 175

Lys 173 Asp 174 Lys 175 Arg 176 Ala Val 177 Ser 178 Gly Ser 179 Val Leu 180  
195 200 205

ASL 210 Adu Gru Eys Gv Ser 1991 Ser Eca Gv The City City City Eys  
210 215 220

225 230 235 240

Aug 11.5 11.5 8.9 245 Aug 11.5 8.9 8.9 8.9 8.9 8.9 250 Aug 11.5 8.9 8.9 8.9 8.9 8.9 255

260 265 270

275                    280                    285  
50                    50                    50

290                    295                    300

305                    310                    315                    320

325 330 335

## EP 1 790 660 A2

Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys  
 340 345 350  
 5 Leu Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp  
 355 360 365  
 Ala Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn  
 370 375 380  
 10 Ile Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp  
 385 390 395 400  
 Glu Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala  
 405 410 415  
 15 Phe Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp  
 420 425 430  
 Glu Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr  
 435 440 445  
 20 Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly  
 450 455 460  
 Lys Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala  
 465 470 475 480  
 25 Glu Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr  
 485 490 495  
 Asn Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr  
 500 505 510  
 30 Arg Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala  
 515 520 525  
 Thr Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile  
 530 535 540  
 35 Ala Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp  
 545 550 555 560  
 Leu Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser  
 565 570 575  
 40 Gly Leu Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala  
 580 585 590  
 Val Gly Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe  
 595 600 605  
 45 Arg Phe Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr  
 610 615 620  
 Ser Ser Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp  
 625 630 635 640  
 50 Leu Glu His His His His His His  
 645  
 <210> 110  
 <211> 1782  
 <212> DNA  
 <213> Artificial Sequence  
 55 <220>  
 <223> deltaG741-961c



	145	150	155	160
5	Thr Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Ile Glu 165 170 175			
	His Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala Ala Asp Ile 180 185 190			
10	Lys Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser Val Leu Tyr 195 200 205			
	Asn Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe Gly Gly Lys 210 215 220			
15	Ala Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val Asn Gly Ile 225 230 235 240			
	Arg His Ile Gly Leu Ala Ala Lys Gln Leu Glu Gly Gly Gly Thr 245 250 255			
20	Gly Ser Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala 260 265 270			
	Ile Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala 275 280 285			
	Gly Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys 290 295 300			
25	Asp Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly 305 310 315 320			
	Leu Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys 325 330 335			
30	Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys 340 345 350			
	Leu Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp 355 360 365			
35	Ala Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn 370 375 380			
	Ile Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp 385 390 395 400			
40	Glu Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala 405 410 415			
	Phe Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp 420 425 430			
45	Glu Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr 435 440 445			
	Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly 450 455 460			
50	Lys Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala 465 470 475 480			
	Glu Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr 485 490 495			
55	Asn Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr 500 505 510			
	Arg Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala			

515

520

525

5           Thr Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile  
      530                           535                           540

Ala Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp  
      545                           550                           555                           560

Leu Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser  
      565                           570                           575

10           Gly Leu Phe Gln Pro Tyr Asn Val Gly Leu Glu His His His His  
      580                           585                           590

His

15

<210> 112  
<211> 3939  
<212> DNA  
<213> Artificial Sequence

20

<220>  
<223> deltaG741-983

<400> 112	atgggtcgccg ccgacatcg tgcgggctt gccgatgcac taaccgcacc gctcgaccat	60
	aaagacaaag gtttgcagt tttgacgtg gatcgtccg tcaggaaaaact cgagaaaactg	120
25	aagctggcg cacaagggtgc gggaaaaact tatggaaacg gtgacagcct caatacgggc	180
	aaattgaaga acgacaagggt cagccgttc gactttatcc gccaaatcga agtggacggg	240
	cagctatta ctttggagag tggagagttc caagtataca aacaaagcca ttccgcctta	300
	accgccttc agaccgagca aatacaagat tcggagcatt ccgggaagat ggttgcgaaa	360
	cgccaggta caatcggcga catagcgggc gaacatacat ctttgaccaa gcttcccga	420
	ggcggcaggg cgacatatcg cgggacggc ttccgttcag acgatgccgg cgggaaaactg	480
30	acctacacca tagatttgcg cgccaaacg gggaaacggca aaatcgaaca ttggaaatcg	540
	ccagaactca atgtcgaccc ggccgcgc gatatacgc cggatggaaa acggccatgcc	600
	gtcatcagcg gttccgtcct ttacaaccaa gccgagaaaag cgacttactc cctcggatc	660
	tttggcgaa aagcccagga aatggcgccg agcgcggaaag tgaaaaccgt aaacggcata	720
	cgccatatcg gccttgcgc caagcaactc gagggatccg gcggaggcgg cacttctgcg	780
	cccgacttca atgcaggcg taccggatc ggcagcaaca gcagagcaac aacagcgaaa	840
	tcagcagcag tatcttacgc cggatatacg aacaaatgt gcaaaagacag aacatgctc	900
35	tgtggccgtc gggatgcgt tgcggttaca gacagggtatc cccaaatcaa tgccccccccc	960
	ccgaaatctgc ataccggaga ctttccaaac ccaaatcgc catacaagaa ttgtatcaac	1020
	ctcaaaccctg caattgaagc aggctataca ggacgcgggg tagaggtagg tattctgcac	1080
	acaggcgaat ccgtcggcag catatccctt cccgaactgt atggcagaaa agaacacggc	1140
	tataacgaaa attacaaaaaa ctataccggc tatatgcggg aggaagcgc tgaagacgga	1200
	ggcggtaaag acattgaagc ttctttcgac gatggggccg ttatagagac tgaagcaaag	1260
40	ccgacggata tccggcacgt aaaagaatc ggacacatcg attttgtctc ccatattatt	1320
	ggcgggctt ccgtggacgg cagacctgc ggcggatttc gcccggatgc gacgctacac	1380
	ataatgaata cgaatgatga aaccaagaac gaaatgatgg ttgcagccat ccgcaatgca	1440
	tgggtcaagc tggcgaacg tggcgtgcg atcgtcaata acagttttgg aacaacatcg	1500
	agggcaggca ctggcgacct ttccaaata gccaattcgg aggagcgtt ccgccaagcg	1560
	ttgctcgact attccggcg tgataaaaaa gacgagggtt tccgcctgt gcaacagagc	1620
	gattacggca acctgtccta ccacatcgta aaaaaaaaaa tgctttcat ctttgcaca	1680
45	ggcaatgacg cacaaggctca gccccacaca tatgccttat tgccatittt tgaaaaagac	1740
	gctcaaaaag gcattatcac agtgcggcgt gtagaccgca gtggagaaaaa gttcaaacgg	1800
	gaaatgtatg gagaaccggg tacagaaccg cttagttagt gctccaacca ttgcggaaatt	1860
	actgccatgt ggtgcctgtc ggcacccatg gaagcaagcg tccgtttcac ccgtacaaac	1920
	ccgattcaaa ttggcggaac atcctttcc gcacccatcg taaccggcac ggcggctctg	1980
	ctgctcgaga aataccctgt gatggacaaac gacaacctgc gtaccacgtt gctgacgacg	2040
50	gctcaggaca tcggtgcagt cggcgtggac agcaaggttc gctggggact gctggatgcg	2100
	ggtaaggcca tgaacggacc cgcgtcttt ccgttgcggc actttaccgc cgatacgaaa	2160
	ggtacatccg atattgccta ctccttccgt aacgacattt caggcacggg cggcctgatc	2220
	aaaaaaggcg gcagccaact gcaactgcac ggcacacaaca cctatacggg caaaaccatt	2280
	atcgaaggcg gttcgctggt gttgtacggc aacaaacaaat cgatatacg cgtcgaaacc	2340
	aaagggtgcgc tgatttataa cggggcggca tccggcggca gcctgaacag cgacggcatt	2400
55	gtctatctgg cagataccga ccaatccggc gcaacacaaa cctgacatcat caaaggcagt	2460
	ctgcagctgg acggcaaagg tacgtgtac acacgtttgg gcaaaactgct gaaagtggac	2520
	ggtacggcga ttatcgccgg caagctgtac atgtcgac gcgcaaggg ggcaggctat	2580

	ctcaacagta	ccggacgacg	tgttcccttc	ctgagtgccg	ccaaaatcg	gcaggattat	2640
5	tctttcttca	caaacatcg	aaccgacggc	ggccctgctgg	tttccctcg	cagcgtcgaa	2700
	aaaacagcgg	gcagtgaagg	cgacacgctg	tcctattatg	tccgtcgcg	aatgcggca	2760
	cggactgctt	cgcgacgccc	acattcccg	cccggccggc	tgaaacacgc	cgtagaacag	2820
	ggcggcagca	atctggaaaa	cctgatggc	gaactggatg	cctccgaaatc	atccgcaaca	2880
	cccgagacgg	ttgaaaactgc	ggcagccgc	cgcacagata	tgccgggcat	ccggccctac	2940
	ggcgcaactt	tccgcgcagc	ggcagccgt	cagcatgcg	atgcccgg	cggtgtacgc	3000
10	atcttcaaca	gtctcgccgc	taccgtctat	gccgacagta	ccgcccggca	tgcgatatg	3060
	cagggacgccc	gcctgaaagc	cgtatcgac	gggttggacc	acaacggcac	gggtctgcgc	3120
	gtcatcgccgc	aaacccaaca	ggacgggtgg	acgtgggaac	agggcggtgt	tgaaggcaaa	3180
	atgcggcgc	gtacccaaac	cgtcgccatt	ggccgcaaaa	ccggcgaaaa	tacgacagca	3240
	gcccccacac	tgggcattggg	acgcagcaca	tggagcgaaa	acagtgc	aaaacc	3300
	gacagcatta	gtctgttgc	aggcatacgg	cacgatgcgg	gcatatcg	ctatctcaaa	3360
15	ggcctgttct	cctacggacg	ctacaaaaac	agcatcagcc	gcagcaccgg	tgcggacgaa	3420
	catgcggaag	gcagcgtcaa	cggcacgctg	atgcagctgg	gcmcactggg	cggtgtcaac	3480
	gttccgttgc	ccgcaacggg	agatttgcg	gtcgaaaggcg	gtctgcgcta	cgacctgctc	3540
	aaacaggatg	cattcgccga	aaaaggcagt	gcttgggct	ggagcggcaa	cagcctcact	3600
	gaaggcacgc	tggtcggact	cgcgggtctg	aagctgtcgc	aaccctttag	cgataaagcc	3660
	gtcctgttgc	caacggcggg	cgtggaacgc	gaccctgaa	gacgcgacta	cacggtaacg	3720
	ggcggcttta	ccggcgcgac	tgcagcaacc	ggcaagacgg	gggcacgcaa	tatgccgcac	3780
20	acccgtctgg	ttgcccggct	gggcgcggat	gtcgaattcg	gcaacggctg	gaacggcttg	3840
	gcacgttaca	gctacgcccgg	ttccaaacag	tacggcaacc	acagcggacg	agtccggcgt	3900
	ggctaccgg	tcctcgagca	ccaccaccac	caccactga			3939
	<210>	113					
	<211>	1312					
	<212>	PRT					
	<213>	Artificial Sequence					
25	<220>						
	<223>	deltaG741-983					
	<400>	113					
	Met Val	Ala Ala Asp Ile Gly Ala Gly Leu	Ala Asp Ala Leu Thr Ala				
30	1	5	10	15			
	Pro Leu Asp His Lys Asp Lys Gly	Leu Gln Ser Leu Thr Leu Asp Gln					
	20	25	30				
	Ser Val Arg Lys Asn Glu Lys	Leu Lys Leu Ala Ala Gln Gly Ala Glu					
35	35	40	45				
	Lys Thr Tyr Gly Asn Gly Asp Ser	Leu Asn Thr Gly Lys Leu Lys Asn					
	50	55	60				
	Asp Lys Val Ser Arg Phe Asp Phe Ile Arg Gln	Ile Glu Val Asp Gly					
40	65	70	75	80			
	Gln Leu Ile Thr Leu Glu Ser Gly Glu	Phe Gln Val Tyr Lys Gln Ser					
	85	90	95				
	His Ser Ala Leu Thr Ala Phe Gln Thr	Glu Gln Ile Gln Asp Ser Glu					
45	100	105	110				
	His Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile Gly Asp Ile						
	115	120	125				
	Ala Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala						
50	130	135	140				
	Thr Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu						
	145	150	155	160			
	Thr Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Ile Glu						
	165	170	175				
55	His Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala Asp Ile						
	180	185	190				

Lys Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser Val Leu Tyr  
 195 200 205  
 5 Asn Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe Gly Gly Lys  
 210 215 220  
 Ala Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val Asn Gly Ile  
 225 230 235 240  
 10 Arg His Ile Gly Leu Ala Ala Lys Gln Leu Glu Gly Ser Gly Gly  
 245 250 255  
 Gly Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser  
 260 265 270  
 15 Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly  
 275 280 285  
 Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg  
 290 295 300  
 20 Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro  
 305 310 315 320  
 Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys  
 325 330 335  
 25 Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg  
 340 345 350  
 Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile  
 355 360 365  
 30 Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn  
 370 375 380  
 Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly  
 385 390 395 400  
 Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu  
 405 410 415  
 35 Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His  
 420 425 430  
 Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg  
 435 440 445  
 40 Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr  
 450 455 460  
 Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala  
 465 470 475 480  
 45 Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe  
 485 490 495  
 Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn  
 500 505 510  
 50 Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp  
 515 520 525  
 Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn  
 530 535 540  
 55 Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr  
 545 550 555 560

## EP 1 790 660 A2

Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe  
 565 570 575  
 5 Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp  
 580 585 590  
 Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr  
 595 600 605  
 10 Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp  
 610 615 620  
 Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn  
 625 630 635 640  
 15 Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly  
 645 650 655  
 Thr Ala Ala Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn Asp Asn  
 660 665 670  
 20 Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly  
 675 680 685  
 Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys Ala Met  
 690 695 700  
 25 Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp Thr Lys  
 705 710 715 720  
 Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser Gly Thr  
 725 730 735  
 30 Gly Gly Leu Ile Lys Lys Gly Ser Gln Leu Gln Leu His Gly Asn  
 740 745 750  
 Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Ser Leu Val Leu  
 755 760 765  
 35 Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly Ala Leu  
 770 775 780  
 Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp Gly Ile  
 785 790 795 800  
 Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His  
 805 810 815  
 40 Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg  
 820 825 830  
 Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys  
 835 840 845  
 45 Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr  
 850 855 860  
 Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr  
 865 870 875 880  
 50 Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala Ser Leu  
 885 890 895  
 Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu Ser Tyr  
 900 905 910  
 55 Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala Ala His  
 915 920 925

Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln Gly Gly Ser Asn  
 930 935 940  
 5 Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu Ser Ser Ala Thr  
 945 950 955 960  
 Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr Asp Met Pro Gly  
 965 970 975  
 10 Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Ala Val Gln His  
 980 985 990  
 Ala Asn Ala Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala Ala Thr  
 995 1000 1005  
 15 Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly Arg Arg  
 1010 1015 1020  
 Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly Leu Arg  
 1025 1030 1035 1040  
 20 Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln Gly Gly  
 1045 1050 1055  
 Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala  
 1060 1065 1070  
 25 Lys Thr Gly Glu Asn Thr Ala Ala Ala Thr Leu Gly Met Gly Arg  
 1075 1080 1085  
 Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser  
 1090 1095 1100  
 30 Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys  
 1105 1110 1115 1120  
 Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr  
 1125 1130 1135  
 35 Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln  
 1140 1145 1150  
 Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp  
 1155 1160 1165  
 40 Leu Thr Val Glu Gly Leu Arg Tyr Asp Leu Leu Lys Gln Asp Ala  
 1170 1175 1180  
 Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser Leu Thr  
 1185 1190 1195 1200  
 45 Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln Pro Leu  
 1205 1210 1215  
 Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val Glu Arg Asp Leu  
 1220 1225 1230  
 50 Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly Ala Thr Ala  
 1235 1240 1245  
 Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His Thr Arg Leu Val  
 1250 1255 1260  
 55 Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly Trp Asn Gly Leu  
 1265 1270 1275 1280  
 Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn His Ser Gly  
 1285 1290 1295

Arg Val Gly Val Gly Tyr Arg Phe Leu Glu His His His His His His  
 1300 1305 1310

5       <210> 114  
 <211> 2028  
 <212> DNA  
 <213> Artificial Sequence

10      <220>  
 <223> deltaG741-ORF46.1

<400> 114  
 atggtcggc cgacatcg tgcgggctt gccatgcac taaccgcacc gctcgaccat 60  
 aaagacaag gttgcagtc ttgcacgtg gatcagtccg tcaggaaaaa cgagaaactg 120  
 aagctggcg cacaagggtc ggaaaaact tatggaaacg gtgacagcct caatacggc 180  
 aaattgaaga acgacaagggt cagccgttc gacttatcc gccaaatcga agtggacggg 240  
 cagtcatta cctggagag tggagagttc caagtataca aacaaagcca ttccgccta 300  
 accgccttc agaccggca aatacaagat tcggagcatt ccggaaagat gttgcgaaa 360  
 cgccaggta gaatcggcga catagcggc gaacatacat ctttgacaa gcttcccga 420  
 ggcggcaggg cgacatatcg cgggacggcg ttcggttcag acgatgcgg cgaaaaactg 480  
 acctacacca tagatttcgc cgccaagcag ggaaacggca aaatcgaaca ttgaaatcg 540  
 15      ccagaactca atgtcgaacct ggccgccc gatatcaagc cggatggaaa acggcatgcc 600  
 gtcatcagc gttccgtctt ttacaaccaa gccgagaaag gcagttactc cctcggtatc 660  
 tttggcgaa aagccaggaa agttgcggc agcggcggaa tgaaaacgt aaacggcata 720  
 cgccatcatcg gccttgcgc caagcaactc gagcgtggcg gaggactgg atccctcagat 780  
 ttggcaaacg attcttttat ccggcagggtt ctgcaccgtc agcatttcga acccgacggg 840  
 20      aaataccacc tattcggcag cagggggaa ctggccgagc gcagcggcca tattggattg 900  
 ggaaaaatac aaagccatca gttggcaac ctgtatgattc aacaggcggc cattaaagga 960  
 aatatcggt acattgtccg ctttccgtat cacggcacg aagtccatc ccccttcgac 1020  
 aaccatcgct cacattccg ttctgtatgaa gccgttagtc cggtgacgg atttagcctt 1080  
 taccgcattcc attgggacgg atacgaacac catccccggc acggctatga cggccacag 1140  
 25      ggcggcggct atcccgctcc caaaggcgcg aggatataat acagctacga cattaaaggc 1200  
 gttgcccata atatccgcct caacctgacc gacaaccgcg gcaccggaca acggcttgcc 1260  
 gaccgttcc acaatgccc tagtatgtc acgcaaggag taggcgacgg attcaaacgc 1320  
 gcccaccat acagccccga gctggacaga tcgggcaatg cccggcgaagc cttaacacggc 1380  
 30      actgcagata tcgtaaaaaa catcatcgcc gggcaggag aaattgtcgg cgcaggcgat 1440  
 gccgtgcagg gataagcga aggtcaaac attgtgtca tgcacggctt ggtctgctt 1500  
 tccaccggaa acaagatggc ggcgcataac gatttggcag atatggcga actcaaagac 1560  
 tatggccgag cagccatccg cgattggca gtccaaaacc ccaatccgc acaaggcata 1620  
 gaagccgtca gcaatatctt tatggcagcc atccccatca aagggattgg agctgttcgg 1680  
 35      ggaaaaatac gcttggcgcatcacggca catcctatca agcggtcgca gatgggcgcg 1740  
 atcgcatatgc cgaaaggaa atccgcgtc agcacaatt ttgccgtatgc ggcatacggc 1800  
 aaataccgt ccccttacca ttccggaaat atccgtcaa acttggcga gcttacggc 1860  
 aaagaaaaca tcacccctc aaccgtgcg ccgtcaaacg gcaaaaatgt caaactggca 1920  
 gaccaacgccc acccgaaagac aggcgttaccg ttgcacggta aagggttcc gaattttgag 1980  
 aagcacgtga aatatgatac gtcgagcac caccaccacc accactga 2028

40      <210> 115  
 <211> 675  
 <212> PRT  
 <213> Artificial Sequence

45      <220>  
 <223> deltaG741-ORF46.1

<400> 115  
 Met Val Ala Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala  
 1                   5                   10                   15

50      Pro Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln  
 20                   25                   30

Ser Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln Gly Ala Glu  
 35                   40                   45

55      Lys Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys Leu Lys Asn  
 50                   55                   60

Asp Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu Val Asp Gly  
 65 70 75 80  
 5 Gln Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr Lys Gln Ser  
 85 90 95  
 His Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln Asp Ser Glu  
 100 105 110  
 10 His Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile Gly Asp Ile  
 115 120 125  
 Ala Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala  
 130 135 140  
 15 Thr Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu  
 145 150 155 160  
 Thr Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Ile Glu  
 165 170 175  
 His Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala Ala Asp Ile  
 180 185 190  
 20 Lys Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser Val Leu Tyr  
 195 200 205  
 25 Asn Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe Gly Gly Lys  
 210 215 220  
 Ala Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val Asn Gly Ile  
 225 230 235 240  
 Arg His Ile Gly Leu Ala Ala Lys Gln Leu Asp Gly Gly Gly Thr  
 245 250 255  
 30 Gly Ser Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp  
 260 265 270  
 Arg Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg  
 275 280 285  
 35 Gly Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln  
 290 295 300  
 Ser His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly  
 305 310 315 320  
 40 Asn Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu Val His  
 325 330 335  
 Ser Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly  
 340 345 350  
 45 Ser Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr  
 355 360 365  
 Glu His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Gly Tyr  
 370 375 380  
 50 Pro Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly  
 385 390 395 400  
 Val Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly  
 405 410 415  
 55 Gln Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln  
 420 425 430

Gly Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu  
 435 440 445  
 5 Asp Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile  
 450 455 460  
 Val Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp  
 465 470 475 480  
 10 Ala Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly  
 485 490 495  
 Leu Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu  
 500 505 510  
 15 Ala Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ala Ile Arg Asp  
 515 520 525  
 Trp Ala Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser  
 530 535 540  
 20 Asn Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg  
 545 550 555 560  
 Gly Lys Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser  
 565 570 575  
 25 Gln Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp  
 580 585 590  
 Asn Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser  
 595 600 605  
 30 Arg Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile  
 610 615 620  
 Thr Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala  
 625 630 635 640  
 35 Asp Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe  
 645 650 655  
 Pro Asn Phe Glu Lys His Val Lys Tyr Asp Thr Leu Glu His His His  
 660 665 670  
 His His His  
 675  
 40 <210> 116  
 <211> 249  
 <212> PRT  
 <213> Artificial Sequence  
 45 <220>  
 <223> Novel protein  
 <400> 116  
 Met Lys Lys Tyr Leu Phe Arg Ala Ala Leu Tyr Gly Ile Ala Ala Ala  
 1 5 10 15  
 50 Ile Leu Ala Ala Ala Ile Pro Ala Gly Asn Asp Ala Thr Thr Lys Pro  
 20 25 30  
 Asp Leu Tyr Tyr Leu Lys Asn Glu Gln Ala Ile Asp Ser Leu Lys Leu  
 35 40 45  
 55 Leu Pro Pro Pro Pro Glu Val Gly Ser Ile Gln Phe Leu Asn Asp Gln  
 50 55 60

## EP 1 790 660 A2

	Ala Met Tyr Glu Lys Gly Arg Met Leu Arg Asn Thr Glu Arg Gly Lys	
5	65 70 75 80	
	Gln Ala Gln Ala Asp Ala Asp Leu Ala Ala Gly Gly Val Ala Thr Ala	
	85 90 95	
	Phe Ser Gly Ala Phe Gly Tyr Pro Ile Thr Glu Lys Asp Ser Pro Glu	
10	100 105 110	
	Leu Tyr Lys Leu Leu Thr Asn Met Ile Glu Asp Ala Gly Asp Leu Ala	
	115 120 125	
	Thr Arg Ser Ala Lys Glu His Tyr Met Arg Ile Arg Pro Phe Ala Phe	
15	130 135 140	
	Tyr Gly Thr Glu Thr Cys Asn Thr Lys Asp Gln Lys Lys Leu Ser Thr	
	145 150 155 160	
	Asn Gly Ser Tyr Pro Ser Gly His Thr Ser Ile Gly Trp Ala Thr Ala	
	165 170 175	
20	Leu Val Leu Ala Glu Val Asn Pro Ala Asn Gln Asp Ala Ile Leu Glu	
	180 185 190	
	Arg Gly Tyr Gln Leu Gly Gln Ser Arg Val Ile Cys Gly Tyr His Trp	
	195 200 205	
25	Gln Ser Asp Val Asp Ala Ala Arg Ile Val Gly Ser Ala Ala Val Ala	
	210 215 220	
	Thr Leu His Ser Asp Pro Ala Phe Gln Ala Gln Leu Ala Lys Ala Lys	
	225 230 235 240	
	Gln Glu Phe Ala Gln Lys Ser Gln Lys	
30	245	
	<210> 117	
	<211> 66	
	<212> DNA	
	<213> Artificial sequence	
35	<220>	
	<223> L1 linker	
	<220>	
	<221> N	
40	<222> 13	
	<223> A, T/U, G or C	
	<400> 117	
	tatgaartay ytnttymgcg ccgcctgtta cggcatcgcc gccgccatcc tcgcccgcgc 60	
	gatccc 66	
45	<210> 118	
	<211> 69	
	<212> DNA	
	<213> Artificial sequence	
	<220>	
50	<223> S1 linker	
	<220>	
	<221> N	
	<222> 25, 28	
	<223> A, T/U, G or C	
55	<400> 118	
	tatgaaaaaa tacctattcc grgcngcnyt rtayggsatc gccgcccaca tcctcgccgc 60	
	cgcgatccc 69	

5           <210> 119  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> 9L1-a  
 <400> 119  
 atgaagaagt accttttcag cgccgcc 27  
 10          <210> 120  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> 9L1-e  
 <400> 120  
 atgaaaaaat acttttccg cgccgcc 27  
 15          <210> 121  
 <211> 27  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> 9L1-d  
 <400> 121  
 atgaaaaaat acttttccg cgccgcc 27  
 20          <210> 122  
 <211> 60  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> 9L1-f  
 <400> 122  
 atgaaaaaat atctcttag cgccgcctg tacggcatcg ccggccat cctcgccgcc 60  
 25          <210> 123  
 <211> 60  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> 919sp  
 <400> 123  
 atgaaaaaat acctattccg cgccgcctg tacggcatcg ccggccat cctcgccgcc 60  
 30          <210> 124  
 <211> 9  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 9L1a  
 <400> 124  
 Met Lys Lys Tyr Leu Phe Ser Ala Ala  
 35          1                 5  
 <210> 125

5           <211> 9  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 9L1e  
 10          <400> 125  
 Met Lys Lys Tyr Phe Phe Arg Ala Ala  
 1            5  
 <210> 126  
 <211> 9  
 <212> PRT  
 <213> Artificial Sequence  
 15          <220>  
 <223> 9L1d  
 <400> 126  
 Met Lys Lys Tyr Phe Phe Arg Ala Ala  
 1            5  
 20          <210> 127  
 <211> 20  
 <212> PRT  
 <213> Artificial Sequence  
 25          <220>  
 <223> 9L1f  
 <400> 127  
 Met Lys Lys Tyr Leu Phe Ser Ala Ala Leu Tyr Gly Ile Ala Ala Ala  
 1            5                   10                   15  
 30          Ile Leu Ala Ala  
 20  
 <210> 128  
 <211> 20  
 <212> PRT  
 35          <213> Artificial Sequence  
 <220>  
 <223> 9L1sp  
 <400> 128  
 Met Lys Lys Tyr Leu Phe Arg Ala Ala Leu Tyr Gly Ile Ala Ala Ala  
 1            5                   10                   15  
 Ile Leu Ala Ala  
 20  
 45          <210> 129  
 <211> 42  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 50          <223> 9S1-e  
 <400> 129  
 atgaaaaaat acctattcat cgccgccc atccctgccg cc                   42  
 <210> 130  
 <211> 60  
 <212> DNA  
 55          <213> Artificial Sequence

5           <220> 130  
 <223> 9S1-c  
 atgaaaaaat acctattccg agctgccaa tacggcatcg ccgcccatt cctcgccgccc 60  
 <210> 131  
 <211> 60  
 <212> DNA  
 <213> Artificial Sequence  
 10          <220> 131  
 <223> 9S1-b  
 <400> 131  
 atgaaaaaat acctattccg ggccgccaa tacggcatcg ccgcccatt cctcgccgccc 60  
 15          <210> 132  
 <211> 60  
 <212> DNA  
 <213> Artificial Sequence  
 20          <220> 132  
 <223> 9S1-i  
 <400> 132  
 atgaaaaaat acctattccg ggcggcttg tacggatcg ccgcccatt cctcgccgccc 60  
 25          <210> 133  
 <211> 14  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 9S1e  
 30          <400> 133  
 Met Lys Lys Tyr Leu Phe Ile Ala Ala Ala Ile Leu Ala Ala  
 1                         5                         10  
 <210> 134  
 <211> 20  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 9S1c  
 40          <400> 134  
 Met Lys Lys Tyr Leu Phe Arg Ala Ala Gln Tyr Gly Ile Ala Ala  
 1                         5                         10                         15  
 Ile Leu Ala Ala  
 45          20  
 <210> 135  
 <211> 20  
 <212> PRT  
 <213> Artificial Sequence  
 50          <220>  
 <223> 9S1b  
 <400> 135  
 Met Lys Lys Tyr Leu Phe Arg Ala Ala Gln Tyr Gly Ile Ala Ala  
 1                         5                         10                         15  
 Ile Leu Ala Ala  
 55          20

5           <210> 136  
 <211> 20  
 <212> PRT  
 <213> Artificial Sequence  
  
 10          <220>  
 <223> 9s1i  
  
 <400> 136  
 Met Lys Lys Tyr Leu Phe Arg Ala Ala Leu Tyr Gly Ile Ala Ala Ala  
 1               5              10               15  
  
 Ile Leu Ala Ala  
 20  
  
 15          <210> 137  
 <211> 467  
 <212> PRT  
 <213> Artificial Sequence  
  
 20          <220>  
 <223> 730  
  
 <400> 137  
 Val Lys Pro Leu Arg Arg Leu Thr Asn Leu Leu Ala Ala Cys Ala Val  
 1               5              10               15  
  
 Ala Ala Ala Ala Leu Ile Gln Pro Ala Leu Ala Ala Asp Leu Ala Gln  
 25            20              25              30  
  
 Asp Pro Phe Ile Thr Asp Asn Ala Gln Arg Gln His Tyr Glu Pro Gly  
 35            35              40              45  
  
 Gly Lys Tyr His Leu Phe Gly Asp Pro Arg Gly Ser Val Ser Asp Arg  
 30            50              55              60  
  
 Thr Gly Lys Ile Asn Val Ile Gln Asp Tyr Thr His Gln Met Gly Asn  
 35            65              70              75              80  
  
 Leu Leu Ile Gln Gln Ala Asn Ile Asn Gly Thr Ile Gly Tyr His Thr  
 40            85              90              95  
  
 Arg Phe Ser Gly His Gly His Glu Glu His Ala Pro Phe Asp Asn His  
 45            100              105              110  
  
 Ala Ala Asp Ser Ala Ser Glu Glu Lys Gly Asn Val Asp Glu Gly Phe  
 50            115              120              125  
  
 Thr Val Tyr Arg Leu Asn Trp Glu Gly His Glu His His Pro Ala Asp  
 55            130              135              140  
  
 Ala Tyr Asp Gly Pro Lys Gly Gly Asn Tyr Pro Lys Pro Thr Gly Ala  
 60            145              150              155              160  
  
 Arg Asp Glu Tyr Thr Tyr His Val Asn Gly Thr Ala Arg Ser Ile Lys  
 65            165              170              175  
  
 Leu Asn Pro Thr Asp Thr Arg Ser Ile Arg Gln Arg Ile Ser Asp Asn  
 70            180              185              190  
  
 Tyr Ser Asn Leu Gly Ser Asn Phe Ser Asp Arg Ala Asp Glu Ala Asn  
 75            195              200              205  
  
 Arg Lys Met Phe Glu His Asn Ala Lys Leu Asp Arg Trp Gly Asn Ser  
 80            210              215              220  
  
 Met Glu Phe Ile Asn Gly Val Ala Ala Gly Ala Leu Asn Pro Phe Ile  
 85            225              230              235              240

## EP 1 790 660 A2

Ser Ala Gly Glu Ala Leu Gly Ile Gly Asp Ile Leu Tyr Gly Thr Arg  
 245 250 255  
 5 Tyr Ala Ile Asp Lys Ala Ala Met Arg Asn Ile Ala Pro Leu Pro Ala  
 260 265 270  
 Glu Gly Lys Phe Ala Val Ile Gly Gly Leu Gly Ser Val Ala Gly Phe  
 275 280 285  
 10 Glu Lys Asn Thr Arg Glu Ala Val Asp Arg Trp Ile Gln Glu Asn Pro  
 290 295 300  
 Asn Ala Ala Glu Thr Val Glu Ala Val Phe Asn Val Ala Ala Ala Ala  
 305 310 315 320  
 15 Lys Val Ala Lys Leu Ala Lys Ala Ala Lys Pro Gly Lys Ala Ala Val  
 325 330 335  
 Ser Gly Asp Phe Ala Asp Ser Tyr Lys Lys Lys Leu Ala Leu Ser Asp  
 340 345 350  
 20 Ser Ala Arg Gln Leu Tyr Gln Asn Ala Lys Tyr Arg Glu Ala Leu Asp  
 355 360 365  
 Ile His Tyr Glu Asp Leu Ile Arg Arg Lys Thr Asp Gly Ser Ser Lys  
 370 375 380  
 25 Phe Ile Asn Gly Arg Glu Ile Asp Ala Val Thr Asn Asp Ala Leu Ile  
 385 390 395 400  
 Gln Ala Lys Arg Thr Ile Ser Ala Ile Asp Lys Pro Lys Asn Phe Leu  
 405 410 415  
 30 Asn Gln Lys Asn Arg Lys Gln Ile Lys Ala Thr Ile Glu Ala Ala Asn  
 420 425 430  
 Gln Gln Gly Lys Arg Ala Glu Phe Trp Phe Lys Tyr Gly Val His Ser  
 435 440 445  
 35 Gln Val Lys Ser Tyr Ile Glu Ser Lys Gly Gly Ile Val Lys Thr Gly  
 450 455 460  
 Leu Gly Asp  
 465  
 40 <210> 138  
 <211> 377  
 <212> PRT  
 <213> Artificial Sequence  
 <220>  
 <223> 730-C1  
 45 <400> 138  
 Met Ala Asp Leu Ala Gln Asp Pro Phe Ile Thr Asp Asn Ala Gln Arg  
 1 5 10 15  
 Gln His Tyr Glu Pro Gly Gly Lys Tyr His Leu Phe Gly Asp Pro Arg  
 20 25 30  
 50 Gly Ser Val Ser Asp Arg Thr Gly Lys Ile Asn Val Ile Gln Asp Tyr  
 35 40 45  
 Thr His Gln Met Gly Asn Leu Leu Ile Gln Gln Ala Asn Ile Asn Gly  
 50 55 60  
 55 Thr Ile Gly Tyr His Thr Arg Phe Ser Gly His Gly His Glu Glu His  
 65 70 75 80

Ala Pro Phe Asp Asn His Ala Ala Asp Ser Ala Ser Glu Glu Lys Gly  
 85 90 95  
 5 Asn Val Asp Glu Gly Phe Thr Val Tyr Arg Leu Asn Trp Glu Gly His  
 100 105 110  
 Glu His His Pro Ala Asp Ala Tyr Asp Gly Pro Lys Gly Gly Asn Tyr  
 115 120 125  
 10 Pro Lys Pro Thr Gly Ala Arg Asp Glu Tyr Thr Tyr His Val Asn Gly  
 130 135 140  
 Thr Ala Arg Ser Ile Lys Leu Asn Pro Thr Asp Thr Arg Ser Ile Arg  
 145 150 155 160  
 15 Gln Arg Ile Ser Asp Asn Tyr Ser Asn Leu Gly Ser Asn Phe Ser Asp  
 165 170 175  
 Arg Ala Asp Glu Ala Asn Arg Lys Met Phe Glu His Asn Ala Lys Leu  
 180 185 190  
 20 Asp Arg Trp Gly Asn Ser Met Glu Phe Ile Asn Gly Val Ala Ala Gly  
 195 200 205  
 Ala Leu Asn Pro Phe Ile Ser Ala Gly Glu Ala Leu Gly Ile Gly Asp  
 210 215 220  
 25 Ile Leu Tyr Gly Thr Arg Tyr Ala Ile Asp Lys Ala Ala Met Arg Asn  
 225 230 235 240  
 Ile Ala Pro Leu Pro Ala Glu Gly Lys Phe Ala Val Ile Gly Gly Leu  
 245 250 255  
 30 Gly Ser Val Ala Gly Phe Glu Lys Asn Thr Arg Glu Ala Val Asp Arg  
 260 265 270  
 Trp Ile Gln Glu Asn Pro Asn Ala Ala Glu Thr Val Glu Ala Val Phe  
 275 280 285  
 Asn Val Ala Ala Ala Ala Lys Val Ala Lys Leu Ala Lys Ala Ala Lys  
 290 295 300  
 35 Pro Gly Lys Ala Ala Val Ser Gly Asp Phe Ala Asp Ser Tyr Lys Lys  
 305 310 315 320  
 Lys Leu Ala Leu Ser Asp Ser Ala Arg Gln Leu Tyr Gln Asn Ala Lys  
 325 330 335  
 40 Tyr Arg Glu Ala Leu Asp Ile His Tyr Glu Asp Leu Ile Arg Arg Lys  
 340 345 350  
 Thr Asp Gly Ser Ser Lys Phe Ile Asn Gly Arg Glu Ile Asp Ala Val  
 355 360 365  
 45 Thr Asn Asp Ala Leu Ile Gln Ala Arg  
 370 375  
 <210> 139  
 <211> 353  
 <212> PRT  
 50 <213> Artificial Sequence  
 <220>  
 <223> 730-c2  
 <400> 139  
 55 Met Ala Asp Leu Ala Gln Asp Pro Phe Ile Thr Asp Asn Ala Gln Arg  
 1 5 10 15

Gln His Tyr Glu Pro Gly Gly Lys Tyr His Leu Phe Gly Asp Pro Arg  
 20 25 30  
 5 Gly Ser Val Ser Asp Arg Thr Gly Lys Ile Asn Val Ile Gln Asp Tyr  
 35 40 45  
 Thr His Gln Met Gly Asn Leu Leu Ile Gln Gln Ala Asn Ile Asn Gly  
 50 55 60  
 10 Thr Ile Gly Tyr His Thr Arg Phe Ser Gly His Gly His Glu Glu His  
 65 70 75 80  
 Ala Pro Phe Asp Asn His Ala Ala Asp Ser Ala Ser Glu Glu Lys Gly  
 85 90 95  
 15 Asn Val Asp Glu Gly Phe Thr Val Tyr Arg Leu Asn Trp Glu Gly His  
 100 105 110  
 Glu His His Pro Ala Asp Ala Tyr Asp Gly Pro Lys Gly Gly Asn Tyr  
 115 120 125  
 20 Pro Lys Pro Thr Gly Ala Arg Asp Glu Tyr Thr Tyr His Val Asn Gly  
 130 135 140  
 Thr Ala Arg Ser Ile Lys Leu Asn Pro Thr Asp Thr Arg Ser Ile Arg  
 145 150 155 160  
 25 Gln Arg Ile Ser Asp Asn Tyr Ser Asn Leu Gly Ser Asn Phe Ser Asp  
 165 170 175  
 Arg Ala Asp Glu Ala Asn Arg Lys Met Phe Glu His Asn Ala Lys Leu  
 180 185 190  
 30 Asp Arg Trp Gly Asn Ser Met Glu Phe Ile Asn Gly Val Ala Ala Gly  
 195 200 205  
 Ala Leu Asn Pro Phe Ile Ser Ala Gly Glu Ala Leu Gly Ile Gly Asp  
 210 215 220  
 Ile Leu Tyr Gly Thr Arg Tyr Ala Ile Asp Lys Ala Ala Met Arg Asn  
 225 230 235 240  
 35 Ile Ala Pro Leu Pro Ala Glu Gly Lys Phe Ala Val Ile Gly Gly Leu  
 245 250 255  
 Gly Ser Val Ala Gly Phe Glu Lys Asn Thr Arg Glu Ala Val Asp Arg  
 260 265 270  
 40 Trp Ile Gln Glu Asn Pro Asn Ala Ala Glu Thr Val Glu Ala Val Phe  
 275 280 285  
 Asn Val Ala Ala Ala Ala Lys Val Ala Lys Leu Ala Lys Ala Ala Lys  
 290 295 300  
 45 Pro Gly Lys Ala Ala Val Ser Gly Asp Phe Ala Asp Ser Tyr Lys Lys  
 305 310 315 320  
 Lys Leu Ala Leu Ser Asp Ser Ala Arg Gln Leu Tyr Gln Asn Ala Lys  
 325 330 335  
 50 Tyr Arg Glu Ala Leu Gly Lys Val Arg Ile Ser Gly Glu Ile Leu Leu  
 340 345 350  
 Gly  
 55 <210> 140  
 <211> 2019

<212> DNA  
 <213> Artificial Sequence  
 5 <220>  
 <223> ORF46.1-741  
 <400> 140  
 atgtcagatt tggcaaacga ttcttttatac cggcaggttc tcgaccgtca gcatttcgaa 60  
 cccgacggga aataccacat attccggcagc agggggggaaac ttggccgagcg cagcggccat 120  
 atcggattgg gaaaaataaca aagccatcg ttgggcAAC tggatgtca acaggcggcc 180  
 attaaaaggaa atatcggtca cattgtccgc ttttccgat acgggcacga agtccattcc 240  
 10 cccttcgaca accatgcctc acattccgat tctgtatgaag ccggtagtcc cggtgacgga 300  
 tttagcctt accgcattca ttgggcggga tacgaaacacc atcccggcga cgctatgac 360  
 gggccacagg gccgcggcta tcccgcctcc aaaggcgcga gggatataa cagctacgac 420  
 ataaaaaggcg ttgccccaaa tatccgcctc aacctgaccg acaaccgcag caccggacaa 480  
 15 cggcttgccg accgttccca caatgcgggt agtatgtga cgcaaggagt aggccgacgga 540  
 ttcaaacgcg ccacccgata cagccccggc ctggacagat cgggcaatgc cgccgaagcc 600  
 ttcaacggca ctgcagatat ctgtaaaaac atatcgccg cggcaggaga aatttgtcggc 660  
 gcaggcgatg ccgtgcaggg cataagcgaa ggctcaaaaca ttgctgtcat gcacggcttg 720  
 ggtctgttt ccaccgaaaaa caagatggcg cgcatcaacg atttggcaga tatggcgc 780  
 ctcaaaagact atgcccgcagc agccatccgc gattgggcag tccaaaaccc caatgccgca 840  
 20 caaggcatag aagccgtcag caatatctt atggcagcca tccccatcaa aggattggg 900  
 gctgttcggg gaaaaatacgg ctggggggc atcacggc acatctatcaa gcggtcgc 960  
 atggggcgcg tcgcatttccca gaaaggggaaa tccggcgtca ggcacaattt tgccgatgcg 1020  
 gcatacgcca aatacccgcc cccatttccat tccggaaata tccgttccaa cttggagcag 1080  
 cgttacggca aagaaaaacat cacctctca accgtgcccgc cgtcaaaacgg caaaaatgtc 1140  
 aaactggcag accaacgcca cccgaagaca ggcgtaccgt ttgacggtaa agggtttccg 1200  
 aattttgaga agcacgtgaa atatgatacg ggatccggag ggggtgggtg cgccgccc 1260  
 25 atcgggtgcgg ggttgcggta tgcaactacc gcaccgcctcg accataaaga caaaggttt 1320  
 cagtcttgcg cgtggatca gtccgtcagg aaaaacgaga aactgaact ggcggcaca 1380  
 ggtgcggaaaaa aaacttatgg aaacgggtac agcctcaata cgggcaattt gaagaacgac 1440  
 aaggctcagcc gtttcgactt tatccggccaa atcgaagtgg acgggcagct cattaccttg 1500  
 gagagtggag agttcaagt atacaaccaa accattccgc cttiaaccgc ctttcagacc 1560  
 gagcaaatac aagattcggg gcattccggg aagatggttg cgaaacgcca gttcagaatc 1620  
 30 ggcgacatag cgggcgaaca tacatctttt gacaagcttc ccgaaggcgg cagggcaca 1680  
 tatcgcggga cggcgttcgg ttccagacat gccggcggaa aactgaccta caccatagat 1740  
 ttcgcggccaa agcaggggaaa cggcaaaatc gaacatttga aatcgccaga actaatgtc 1800  
 gacctggccg cggccgatata caagccggat ggaaaaacgccc atgcccgtcat cagcgggttcc 1860  
 gtcctttaca accaagccga gaaaggcagt tactccctcg gtatctttgg cgaaaaagcc 1920  
 caggaagttg cggcagcgc ggaagtggaaa accgtaaacg gcatacgcca tattcggcctt 1980  
 35 gccgcgaagc aactcgagca ccaccaccac caccactga 2019

<210> 141  
 <211> 672  
 <212> PRT  
 <213> Artificial Sequence

40 <220>  
 <223> ORF46.1-741  
 <400> 141  
 Met Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg  
 1 5 10 15  
 45 Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg Gly  
 20 25 30  
 Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln Ser  
 35 40 45  
 50 His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly Asn  
 50 55 60  
 Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu Val His Ser  
 65 70 75 80  
 55 Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly Ser  
 85 90 95

Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu  
 100 105 110  
 5 His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Tyr Pro  
 115 120 125  
 Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val  
 130 135 140  
 10 Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln  
 145 150 155 160  
 Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly  
 165 170 175  
 15 Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp  
 180 185 190  
 Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val  
 195 200 205  
 20 Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp Ala  
 210 215 220  
 Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu  
 225 230 235 240  
 25 Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala  
 245 250 255  
 Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ile Arg Asp Trp  
 260 265 270  
 30 Ala Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn  
 275 280 285  
 Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly  
 290 295 300  
 35 Lys Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln  
 305 310 315 320  
 Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn  
 325 330 335  
 Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg  
 340 345 350  
 40 Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr  
 355 360 365  
 Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp  
 370 375 380  
 45 Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro  
 385 390 395 400  
 Asn Phe Glu Lys His Val Lys Tyr Asp Thr Gly Ser Gly Gly Gly  
 405 410 415  
 50 Val Ala Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala Pro  
 420 425 430  
 Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln Ser  
 435 440 445  
 55 Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln Gly Ala Glu Lys  
 450 455 460

EP 1 790 660 A2

Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys Leu Lys Asn Asp  
 465 470 475 480  
 Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu Val Asp Gly Gln  
 485 490 495 500  
 Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr Lys Gln Ser His  
 500 505 510 515  
 Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln Asp Ser Glu His  
 515 520 525 530  
 Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile Gly Asp Ile Ala  
 530 535 540 545  
 Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala Thr  
 545 550 555 560  
 Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu Thr  
 565 570 575 580  
 Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Ile Glu His  
 580 585 590 595  
 Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala Asp Ile Lys  
 595 600 605 610  
 Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser Val Leu Tyr Asn  
 610 615 620 625  
 Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe Gly Gly Lys Ala  
 625 630 635 640  
 Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val Asn Gly Ile Arg  
 645 650 655 660  
 His Ile Gly Leu Ala Ala Lys Gln Leu Glu His His His His His His  
 660 665 670

<210> 142  
<211> 2421  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> ORF46.1-961

	aattttgaga	agcacgtcaa	atatgatacg	ggatccggag	gaggaggagc	cacaacgcac	1260
5	gacgatgtta	aaaaagctgc	caactgtggcc	attgctgctg	cctacaacaa	tggcaagaaa	1320
	atcaacccgtt	tcaaagctgg	agagaccatc	tacgacattg	atgaagacgg	cacaattacc	1380
	aaaaaaagacg	caactgcgc	cgatgtgaa	gccgacgact	ttaaaaggct	gggtctgaaa	1440
	aaagtctgaa	ctaaccgtac	caaaaacgtc	aatgaaaaca	aacaaaacgt	cgatgcca	1500
	gtaaaaagctg	cagaatctga	aatagaaaag	ttaacaacca	agtttagcaga	cactgtatgcc	1560
	gcttttagcag	atactgatgc	cgctctggat	gcaaccacca	acgccttgaa	taaattggga	1620
	gaaaatataa	cgacatttg	tgaagagact	aagacaaata	tcgtaaaaat	tgatgaaaaa	1680
10	ttagaagccg	tggtctgatac	cgtcgacaag	catgcccgaag	cattcaacga	tatcgccgat	1740
	tcattggatg	aaaccaacac	taaggcagac	gaagccgtca	aaaccgcca	tgaagccaaa	1800
	cagacggccg	aagaaaccaa	acaaaacgtc	gatgccaaag	taaaaggctc	agaaaactgca	1860
	gcaggcCAAAG	ccgaagctgc	cgctggcaca	gctaatactg	cagccgacaa	ggccgaaagct	1920
	gtcgctgcaa	aagttaccga	catcaaagct	gatatacgta	cgaacaaaga	taatattgct	1980
15	aaaaaaagcaa	acagtgcgc	cgtgtacacc	agagaagagt	ctgacagcaa	atttgcaga	2040
	attgtatggtc	tgaacgctac	taccgaaaaa	ttggacacac	gcttggcttc	tgctgaaaaa	2100
	tccattggcg	atcacgatac	tcgcctgaa	ggtttggata	aaacagtgtc	agacactgcgc	2160
	aaagaaaccc	gccaaggccc	tgcagaaaca	gccgcgcctc	ccggctgtt	ccacacccat	2220
	aacgtgggtc	ggttcaatgt	aacggctgca	gtcggcggct	acaaatccga	atcggcagtc	2280
	gccatcgta	ccggcttccg	ctttaccgaa	aacttgcgc	ccaaagcagg	cgtggcagtc	2340
	ggcacttcgt	ccggcttcc	cgcagcctac	catgtcggcg	tcaattacga	gtggctcgag	2400
	caccaccacc	accaccactg	a				2421
20	<210>	143					
	<211>	806					
	<212>	PRT					
	<213>	Artificial Sequence					
25	<220>						
	<223>	ORF46.1-961					
	<400>	143					
	Met Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg						
	1	5	10		15		
30	Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg Gly						
	20	25	30				
	Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln Ser						
	35	40	45				
35	His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly Asn						
	50	55	60				
	Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu Val His Ser						
	65	70	75		80		
40	Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly Ser						
	85	90	95				
	Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu						
	100	105	110				
45	His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Tyr Pro						
	115	120	125				
	Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val						
	130	135	140				
	Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln						
	145	150	155		160		
50	Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly						
	165	170	175				
	Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp						
	180	185	190				
55	Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val						
	195	200	205				

Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp Ala  
 210 215 220  
 5 Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu  
 225 230 235 240  
 Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala  
 245 250 255  
 10 Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ile Arg Asp Trp  
 260 265 270  
 Ala Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn  
 275 280 285  
 15 Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly  
 290 295 300  
 Lys Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln  
 305 310 315 320  
 20 Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn  
 325 330 335  
 Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg  
 340 345 350  
 25 Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr  
 355 360 365  
 Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp  
 370 375 380  
 30 Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro  
 385 390 395 400  
 Asn Phe Glu Lys His Val Lys Tyr Asp Thr Gly Ser Gly Gly Gly  
 405 410 415  
 35 Ala Thr Asn Asp Asp Val Lys Ala Ala Thr Val Ala Ile Ala  
 420 425 430  
 Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly Glu  
 435 440 445  
 Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp Ala  
 450 455 460  
 40 Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu Lys  
 465 470 475 480  
 Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln Asn  
 485 490 495  
 45 Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu Thr  
 500 505 510  
 Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala Ala  
 515 520 525  
 50 Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile Thr  
 530 535 540  
 Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu Lys  
 545 550 555 560  
 55 Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe Asn  
 565 570 575

Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu Ala  
 580 585 590  
 5 Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys Gln  
 595 600 605  
 Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys Ala  
 610 615 620  
 10 Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu Ala  
 625 630 635 640  
 Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn Lys  
 645 650 655  
 15 Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg Glu  
 660 665 670  
 Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr Thr  
 675 680 685  
 20 Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala Asp  
 690 695 700  
 His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu Arg  
 705 710 715 720  
 25 Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu  
 725 730 735  
 Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val Gly  
 740 745 750  
 30 Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg Phe  
 755 760 765  
 Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser Ser  
 770 775 780  
 35 Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp Leu Glu  
 785 790 795 800  
 His His His His His  
 805  
 40 <210> 144  
 <211> 2256  
 <212> DNA  
 <213> Artificial Sequence  
 <220>  
 <223> ORF46.1-961c  
 45 <400> 144  
 atgtcagatt tggcaaacga ttcttttatac cggcagggttc tcgaccgtca gcatttcgaa 60  
 cccgacggga aataccacct attcggcagc aggggggaac ttgccgagcg cagcggccat 120  
 atcggattgg gaaaaataca aagccatcgat ttgggcaacc ttagtgcattca acaggcggcc 180  
 attaaaggaa atatcggtt cattgtccgc tttccgcatt acgggcacga agtccattcc 240  
 cccttcgaca accatgcctc acattccgat tctgtatggaa cggtagtgc cggtgacgga 300  
 tttagccctt accgcattca ttgggacgga tacgaacacc atcccgcgc cggctatgac 360  
 gggccacagg gcgccggcta tccccctccc aaaggcgcga gggatatac cagctacgac 420  
 ataaaaggcg ttgccccaaa tatccgcctc aaccgtaccg acaaccgcag caccggacaa 480  
 cggcttgccg accgtttcca caatgccgtt agtgcgttgc cgcaggagt aggacggaa 540  
 ttcaaacgcg ccacccgata cagccccggat ctggacagat cggcaatgc cggcgaagcc 600  
 ttcaacgcgca ctgcagat cgttaaaaac atcatcgccg cggcaggaga aattgtcggc 660  
 55 gcaggcgtat ccgtcgaggc cataaggcga ggctcaaaca ttgctgtcat gcacggcttg 720  
 ggtctgtttt ccacccgaaaa caagatggcg cgcacatcaacg atttggcaga tatggcgaa 780  
 ctcacaaagact atgcccgcagc agccatccgc gattggcag tccaaaaccc caatgcccga 840

	caaggcata	aagccgtc	aatatctt	atggcagcc	tccccatcaa	agggattgg	900										
	gctttcgg	gaaaatacgg	cttggcg	atcacggc	atcctatcaa	gcgtcg	960										
5	atggcgcg	tcgcattgc	gaaaggaaa	tccggcgt	gcgacaattt	tgcgatgc	1020										
	gcatacgcc	aatacccgt	cccttacat	tccgaaata	tccgttcaa	cttgagcag	1080										
	cgttacgg	agaaaaacat	caccttc	accgtgc	cgtcaaacgg	aaaaatgtc	1140										
	aaactggc	accacgc	cccgaagaca	ggcttaccgt	ttgacggtaa	agggttccg	1200										
10	aattttgaga	agcacgtaa	atatgat	ggatccggag	gaggaggagc	cacaacgc	1260										
	gacgatgtt	aaaaaagctgc	cactgtggc	attgtctgt	cctacaacaa	tggccaagaa	1320										
	atcaacgg	ttcaaagctgg	agagaccatc	tacgacattt	atgaagacgg	cacaattacc	1380										
	aaaaaaagac	caactgc	cgatgttga	gccgacact	ttaaagg	gggtctgaaa	1440										
	aaagtctg	ctaaccgt	aaaaccgt	aatggaaaaca	aacaaaacgt	cgatgc	1500										
15	gtaaaagctg	cagaatctg	aatagaaaag	ttaacaacca	atgtacgaga	cactgtatgc	1560										
	gctttacg	atactgtatgc	cgcttggat	gcaaccacca	acgcctt	taatttggaa	1620										
	aaaaatataa	cgacatttgc	tgaagagact	aagacaaata	tgc	taatattggaa	1680										
	tttagaagcc	tggctgatac	cgtcgacaag	catgccc	cattcaacga	tatcgccgat	1740										
	tcattggat	aaacccaacac	taaggcagac	gaagccgt	aaacc	tgaagccaaa	1800										
	cagacgccc	aaagaaacccaa	acaaaacgtc	gatgc	taaaagctgc	agaaactgca	1860										
	gcaggc	ccgaaagctgc	cgcttggcaca	gctaataactg	cagccgacaa	ggccgaaagct	1920										
	gtcgtc	aaagtattcga	catcaaagct	gatatcgct	cgaacaaaga	taatattgct	1980										
20	aaaaaaagac	acagtgc	cgtgtacacc	agagaagagt	ctgacagcaa	atttgtcaga	2040										
	attgtatgtt	tgaacgt	taccggaaaa	ttggacacac	gcttgg	tgtgaaaaaa	2100										
	tccatttgc	atcacgatac	tcgcctgt	ggtttggata	aaacagtgtc	agacctgc	2160										
	aaagaaaccc	gccaaggcct	tgcagaacaa	gcccgc	ccgg	tctgtt	2220										
	aacgtgg	tcgagc	ccaccac	cactga		ccaac	2256										
25	<210>	145															
	<211>	751															
	<212>	PRT															
	<213>	Artificial Sequence															
	<220>																
	<223>	ORF46.1-961c															
	<400>	145															
30	Met	Ser	Asp	Leu	Ala	Asn	Asp	Ser	Phe	Ile	Arg	Gln	Val	Leu	Asp	Arg	
	1				5					10				15			
	Gln	His	Phe	Glu	Pro	Asp	Gly	Lys	Tyr	His	Leu	Phe	Gly	Ser	Arg	Gly	
								20			25			30			
35	Glu	Leu	Ala	Glu	Arg	Ser	Gly	His	Ile	Gly	Leu	Gly	Lys	Ile	Gln	Ser	
								35		40			45				
	His	Gln	Leu	Gly	Asn	Leu	Met	Ile	Gln	Gln	Ala	Ala	Ile	Lys	Gly	Asn	
							50		55			60					
40	Ile	Gly	Tyr	Ile	Val	Arg	Phe	Ser	Asp	His	Gly	His	Glu	Val	His	Ser	
							65		70			75			80		
	Pro	Phe	Asp	Asn	His	Ala	Ser	His	Ser	Asp	Ser	Asp	Glu	Ala	Gly	Ser	
							85		90			95					
45	Pro	Val	Asp	Gly	Phe	Ser	Leu	Tyr	Arg	Ile	His	Trp	Asp	Gly	Tyr	Glu	
							100		105			110					
	His	His	Pro	Ala	Asp	Gly	Tyr	Asp	Gly	Pro	Gln	Gly	Gly	Tyr	Pro		
							115		120			125					
50	Ala	Pro	Lys	Gly	Ala	Arg	Asp	Ile	Tyr	Ser	Tyr	Asp	Ile	Lys	Gly	Val	
							130		135			140					
	Ala	Gln	Asn	Ile	Arg	Leu	Asn	Leu	Thr	Asp	Asn	Arg	Ser	Thr	Gly	Gln	
							145		150			155			160		
55	Arg	Leu	Ala	Asp	Arg	Phe	His	Asn	Ala	Gly	Ser	Met	Leu	Thr	Gln	Gly	
							165		170			175					
	Val	Gly	Asp	Gly	Phe	Lys	Arg	Ala	Thr	Arg	Tyr	Ser	Pro	Glu	Leu	Asp	
							180		185			190					

Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val  
 195 200 205  
 5 Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp Ala  
 210 215 220  
 Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu  
 225 230 235 240  
 10 Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala  
 245 250 255  
 Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ile Arg Asp Trp  
 260 265 270  
 15 Ala Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn  
 275 280 285  
 Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly  
 290 295 300  
 20 Lys Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln  
 305 310 315 320  
 Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn  
 325 330 335  
 25 Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg  
 340 345 350  
 Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr  
 355 360 365  
 30 Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp  
 370 375 380  
 Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro  
 385 390 395 400  
 35 Asn Phe Glu Lys His Val Lys Tyr Asp Thr Gly Ser Gly Gly Gly  
 405 410 415  
 Ala Thr Asn Asp Asp Val Lys Ala Ala Thr Val Ala Ile Ala  
 420 425 430  
 40 Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly Glu  
 435 440 445  
 Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp Ala  
 450 455 460  
 45 Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu Lys  
 465 470 475 480  
 Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln Asn  
 485 490 495  
 50 Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu Thr  
 500 505 510  
 Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala Ala  
 515 520 525  
 Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Glu Asn Ile Thr  
 530 535 540  
 55 Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu Lys  
 545 550 555 560

Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe Asn  
 565 570 575  
 5 Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu Ala  
 580 585 590  
 Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys Gln  
 595 600 605  
 10 Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys Ala  
 610 615 620  
 Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu Ala  
 625 630 635 640  
 15 Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn Lys  
 645 650 655  
 Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg Glu  
 660 665 670  
 20 Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr Thr  
 675 680 685  
 Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala Asp  
 690 695 700  
 25 His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu Arg  
 705 710 715 720  
 Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly Leu  
 725 730 735  
 30 Phe Gln Pro Tyr Asn Val Gly Leu Glu His His His His His His  
 740 745 750  
 <210> 146  
 <211> 2421  
 <212> DNA  
 <213> Artificial Sequence  
 35 <220>  
 <223> 961-ORF46.1  
 <400> 146  
 atggccacaa acgacgacga tggtaaaaaa gctgccactg tggccattgc tgctgcctac 60  
 40 aacaatggcc aagaaatcaa cggtttcaaa gctggagaga ccatctacga cattgatgaa 120  
 gacggcacaa ttaccaaaaa agacgcaact gcagccgatg ttgaagccga cgactttaaa 180  
 ggtctgggtc tggaaaaaaa cgtgactaac ctgacccaaa cggtaatga aaacaacaa 240  
 aacgtcgatg ccaaagtaaa agctgcagaa tctgaaatag aaaagttAAC aaccaagtt 300  
 gcagacactg atgcccgttt agcagatact gatgccgctc tggatgcaac caccAACGCC 360  
 ttgaaataat tggagaaaaa tataacgaca tttgctgaag agactaagac aaatatcgta 420  
 aaaattgtatg aaaaatttaga agccgtgct gataccgtcg caacAGATGC cgaAGCATTc 480  
 45 aacgatatacg ccgattcatt ggatgaaacc aacactaagg cagacGAAGC cgtcaAAAC 540  
 gccaatgaag ccaaACAGAC ggCCGAAGAA accAAACAAA acgtcgatgc caaAGTAAAA 600  
 gctgcagaaa ctgcagcagg caaAGCCGAA gctGCCGCTG gcacAGCTAA tactgcagCC 660  
 gacaaggCCG aagtgtcgc tgcaaaAGTT accgacatca aagctgtatc cgctacgaac 720  
 aaagataaaat ttgctaaaaa agcaaaACGT gccgacgtgt acaccAGAGA agagtctgac 780  
 50 agcaaaATTG tcagaattga tggtctgaac gctactaccg AAAAATTGGA cacacgCTTG 840  
 gcttctgctg aaaaatccat tgccgatCAC gataCTCGCC tgaacCGTTT ggataAAACA 900  
 gtgtcagacc tgcgcaaaAGA aACCCGCCA ggcCTTGCAG aacaAGCCG GCTTCGGT 960  
 ctgttccaac cttacaacgt gggTCGGTT aatgtAACGG ctgcAGTCGG cggCTACAAA 1020  
 tccgaatCgg cagtgcGCCat cggTACCGGC ttccgcTTA cggAAAACTT tgccGCCAAA 1080  
 55 gcaggcgtgg cagtgcggcAC ttCGTCCGGT tCTTCCGAG CCTACCATGT cggcgtcaat 1140  
 tacgagtggg gatCCGGAGG aggAGGATCA gatTTGGCAA acgattCTT tatCCGGCAG 1200  
 gttctcgacc gtcaGcATTt cgaACCCGAC gggAAATACC acctATTcGG cagcAGGGGG 1260  
 gaacttgccg agcgcAGCGG ccatATCGGA ttggaaaaaa tacaAAAGCCA tcagTTGGC 1320  
 aacctgtatga ttcaacAGGC ggcattaaaa gggAAATATCG gctacATTGT ccgCTTTCC 1380

5	gatcacgggc acgaagtcca ttcccccttc gacaaccatg cctcacattc cgattctgat gaagccggttta gtcccggttga cggatttagc ctttacgc tccattggga cggatcacaa caccatcccg ccgacggcta tgacggggca caggcgccg gctatcccgc tcccaaaggc gcgagggata tatacagcta cgacataaaa ggcgttgcggc aaaatatccg cctcaacctg accgacaacc gcagcaccgg acaacggctt gccgaccgtt tccacaatgc cggttagtatg ctgacgcaga gacttaggcga cggattcaaa cgcgccaccc gatacagccc cgagctggac agatcgggca atggccgcga agccttcaac ggcactgcag atatcgtaa aaacatcatc ggcgccgcag gagaatttgt cggcgcaggc gatggcgtgc agggcataag cgaaggctca aacattgtc tcatgcacgg ctgggtctg ctggcaccgg aaaacaagat ggccgcgcac aacgatttg cagatatggc gcaactcaaa gactatggc cagcagccat ccgcgattgg gcagtccaaa accccaaatgc cgcacaaggc atagaagccg tcagcaatataat ctttatggca gccatccccca tcaaaggat tggagctgtt cggggaaaat acggcttggg cggcatc gcacatccca tcaagcggcgc gcaatgggc gcgatcgc tggcggaaagg gaaatccgc gtcagcaca attttgcga tgcggcatac gccaatacc cgtcccccta ccattccgc aatatccgtt caaacttggc gcaatggc ggcggaaacatcaccc tcgaccgtg ccgcccgttca acggcaaaaaa tgtcaaactg gcagaccaac gccacccgaa gacaggcgt ccggttgcacg gtaaagggtt tccgaatttt gagaagcacg tgaaatatgta tacgctcgag caccaccacc accaccactg a	1440 1500 1560 1620 1680 1740 1800 1860 1920 1980 2040 2100 2160 2220 2280 2340 2400 2421
10	<210> 147 <211> 806 <212> PRT 20 <213> Artificial Sequence	
20	<220> <223> 961-ORF46.1	
25	<400> 147 Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile 1 5 10 15	
30	Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly 20 25 30	
35	Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp 35 40 45	
40	Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu 50 55 60	
45	Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln 65 70 75 80	
50	Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu 85 90 95	
55	Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala 100 105 110	
60	Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile 115 120 125	
65	Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu 130 135 140	
70	Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe 145 150 155 160	
75	Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu 165 170 175	
80	Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys 180 185 190	
85	Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys 195 200 205	
90	Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu 210 215 220	

Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn  
 225 230 235 240  
 5 Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg  
 245 250 255  
 Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr  
 260 265 270  
 10 Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala  
 275 280 285  
 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu  
 290 295 300  
 15 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly  
 305 310 315 320  
 Leu Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val  
 325 330 335  
 20 Gly Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg  
 340 345 350  
 Phe Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser  
 355 360 365  
 25 Ser Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp Gly  
 370 375 380  
 Ser Gly Gly Gly Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln  
 385 390 395 400  
 30 Val Leu Asp Arg Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe  
 405 410 415  
 Gly Ser Arg Gly Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly  
 420 425 430  
 Lys Ile Gln Ser His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala  
 435 440 445  
 35 Ile Lys Gly Asn Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His  
 450 455 460  
 Glu Val His Ser Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp  
 465 470 475 480  
 40 Glu Ala Gly Ser Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp  
 485 490 495  
 Asp Gly Tyr Glu His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly  
 500 505 510  
 45 Gly Gly Tyr Pro Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp  
 515 520 525  
 Ile Lys Gly Val Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg  
 530 535 540  
 50 Ser Thr Gly Gln Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met  
 545 550 555 560  
 Leu Thr Gln Gly Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser  
 565 570 575  
 55 Pro Glu Leu Asp Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr  
 580 585 590

## EP 1 790 660 A2

Ala Asp Ile Val Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly  
 5 595 600 605  
 Ala Gly Asp Ala Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala val  
 610 615 620  
 Met His Gly Leu Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile  
 625 630 635 640  
 10 Asn Asp Leu Ala Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala  
 645 650 655  
 Ile Arg Asp Trp Ala val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu  
 660 665 670  
 Ala Val Ser Asn Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly  
 15 675 680 685  
 Ala Val Arg Gly Lys Tyr Gly Leu Gly Gly Ile Thr Ala His Pro Ile  
 690 695 700  
 20 Lys Arg Ser Gln Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala  
 705 710 715 720  
 Val Ser Asp Asn Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro  
 725 730 735  
 25 Tyr His Ser Arg Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys  
 740 745 750  
 Glu Asn Ile Thr Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val  
 30 755 760 765  
 Lys Leu Ala Asp Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly  
 770 775 780  
 Lys Gly Phe Pro Asn Phe Glu Lys His val Lys Tyr Asp Thr Leu Glu  
 785 790 795 800  
 His His His His His His  
 35 805  
 <210> 148  
 <211> 1938  
 <212> DNA  
 <213> Artificial Sequence  
 40 <220>  
 <223> 961-741  
 <400> 148  
 atggccacaa acgacgacga tggtaaaaaaa gctgccactg tggccatgc tgctgcctac  
 45 aacaatggcc aagaaaatcaa cggtttcaaa gctggagaga ccatctacga cattgtatgaa  
 gacggcacaa ttaccaaaaa agacgcaact gcagccgatg ttgaagccga cgactttaaa  
 ggtctgggtc tggaaaaagt cgtgactaac ctgaccaaaa ccgtcaatga aaacaaacaa  
 aacgtcgatg ccaaagtaaa agctgcagaa tctgaaatag aaaagttaac aaccaagtt  
 50 gcagacactg atggcgcttt agcagatact gatgccgctc tggatgcac caccaacgccc  
 ttgaataaat tggagaaaaa tataacgaca tttgtctgaaag agactaagac aaatatcgta  
 aaaatttgatg aaaaatttaga agccgtggct gataccgtcg acaaggatgc cgaagcattc  
 aacgatatcg ccgattcatt ggatgaaaacc aacactaagg cagacgaagc cgtcaaaacc  
 55 gccaatgaag ccaaacagac ggccgaagaa accaaacaaa acgtcgatgc caaagtaaaa  
 gctgcagaaaa ctgcagcagg caaagccgaa gctgccgctg gcacagctaa tactgcagcc  
 gacaaggccg aagctgtcgc tgcaaaaagt accgacatca aagctgatat cgctacgaac  
 aaagataata ttgtaaaaaa agcaaacagt gcccacgtgt acaccagaga agagtctgac  
 agcaaatttg tcagaattga tggtctgaac gctactaccg aaaaatttggc cacacgcttg  
 gcttcgtctg aaaaatccat tgccgatcac gatactcgcc tgaacggttt ggataaaaaca  
 60 660 720 780 840 900 960 1020  
 gtgtcagacc tgcgcaaaaga aaccgccaa ggccttgag aacaagccgc gctctccggt  
 65 ctgttccaac cttacaacgt gggtcgggttc aatgtaacgg ctgcagtcgg cggtcacaaa

	tccgaatcg	cagtgc	cggtaccggc	ttccgcttta	ccgaaaactt	tgccgccaaa	1080	
	gcaggcgtgg	cagtgc	cgac	ttcgtccgt	tctccgcag	cctaccatgt	1140	
5	tacgagtggg	gatccggagg	gggtgggtgc	gccgcgcaca	tcgggtgcggg	gcttgcgcgt	1200	
	gcactaaccg	caccgc	tcga	ccataaagac	aaaggttgc	agtctttgc	1260	
	tccgtcagg	aaaaacggaga	actgaagctg	gcggcacaag	gtgcggaaaa	aacttatgg	1320	
	aacggtgaca	gcctcaatac	gggcaatttgc	aagaacgaca	aggtcagccg	tttcgacttt	1380	
	atccgc	aaa	cggcagctc	attaccttgg	agagtggaga	gttccaagta	1440	
	tacaaacaaa	gcattccgc	cttaaccgc	tttcagaccg	agcaaataca	agattcggag	1500	
10	cattccgg	ga	gaaacgc	ttcagaatcg	g	ggcgaacat	1560	
	acatcttt	tttgc	cgaggcggc	agggcgcacat	atcgcgg	ggcg	1620	
	tcagacatg	ccggcg	aaa	actgacatc	accatagatt	tcgccc	gcaggaaac	1680
	ggcaaaatcg	aacattt	gaa	atcgc	ctcaatgtcg	acctgg	cgccgat	1740
	aaggcggatg	gaaaacgc	tgccgtc	atcgc	ttcgttccg	tcctt	ccaagccg	1800
	aaaggcagtt	actcc	ctcg	tatctt	ggaaaagcc	aggaagtgc	cgccagc	1860
	gaagtgaaaa	ccgtaaacgg	catac	ccat	atcggc	ccgcca	actcgagc	1920
	caccaccacc	accactga						1938
15	<210>	149						
	<211>	645						
	<212>	PRT						
	<213>	Artificial Sequence						
20	<220>							
	<223>	961-741						
	<400>	149						
	Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile							
	1	5	10				15	
25	Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly							
	20	25	30					
	Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp							
	35	40	45					
30	Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu							
	50	55	60					
	Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln							
	65	70	75				80	
35	Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu							
	85	90	95					
	Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala							
	100	105	110					
40	Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile							
	115	120	125					
	Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu							
	130	135	140					
45	Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe							
	145	150	155				160	
	Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu							
	165	170	175					
50	Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys							
	180	185	190					
	Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys							
	195	200	205					
55	Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu							
	210	215	220					
	Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn							

## EP 1 790 660 A2

	225	230	235	240
5	Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg 245 250 255			
	Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr 260 265 270			
10	Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala 275 280 285			
	Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu 290 295 300			
15	Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly 305 310 315 320			
	Leu Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val 325 330 335			
20	Gly Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg 340 345 350			
	Phe Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser 355 360 365			
	Ser Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp Gly 370 375 380			
25	Ser Gly Gly Gly Val Ala Ala Asp Ile Gly Ala Gly Leu Ala Asp 385 390 395 400			
	Ala Leu Thr Ala Pro Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu 405 410 415			
30	Thr Leu Asp Gln Ser Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala 420 425 430			
	Gln Gly Ala Glu Lys Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly 435 440 445			
35	Lys Leu Lys Asn Asp Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile 450 455 460			
	Glu Val Asp Gly Gln Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val 465 470 475 480			
40	Tyr Lys Gln Ser His Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile 485 490 495			
	Gln Asp Ser Glu His Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg 500 505 510			
45	Ile Gly Asp Ile Ala Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu 515 520 525			
	Gly Gly Arg Ala Thr Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala 530 535 540			
50	Gly Gly Lys Leu Thr Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn 545 550 555 560			
	Gly Lys Ile Glu His Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala 565 570 575			
55	Ala Ala Asp Ile Lys Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly 580 585 590			
	Ser Val Leu Tyr Asn Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile			

	595	600	605	
5	Phe Gly Gly Lys Ala Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr 610 615 620			
10	Val Asn Gly Ile Arg His Ile Gly Leu Ala Ala Lys Gln Leu Glu His 625 630 635 640			
15	His His His His His 645			
20	<210> 150 <211> 4335 <212> DNA <213> Artificial Sequence			
25	<220> <223> 961-983			
30	<400> 150 atggccacaa acgacgacga tgtaaaaaaa gctgccactg tggccattgc tgctgcctac 60 aacaatggcc aagaaatcaa cggttcaaa gctggagaga ccatctacga cattgatgaa 120 gacggcacaa ttacaaaaaa agacgcaact gcggccgatg ttgaaggcga cgactttaaa 180 ggtctgggtc tggaaaaaaat cgtgactaac ctgaccaaaa ccgtcaatga aaacaaacaa 240 aacgtcgatg ccaaagtaaa agctgcagaa tctgaaaatag aaaagttAAC aaccaagtt 300 gcagacactg atgcccgttt agcagatact gatgccgctc tggatgcaac caccacgccc 360 ttgaataaat tgggagaaaa tataacgaca ttgtctgaag agactaagac aaatatcgta 420 aaaattgtatg aaaaatttaga agccgtggct gataccgtcg acaagcatgc cgaagcattc 480 aacgatatcg cggatttcattt ggtatgaaacc aacactaagg cagacaaagc cgtcaaaacc 540 gccaatgaaag ccaaacagac ggccgaagaa accaaaaaaa acgtcgatgc caaagtaaaa 600 gctgcagaaa ctgcagcagg caaagccgaa gctgccgtc gcacagctaa tactgcagcc 660 gacaaggccg aagctgtcgc tgcaaaagtt accgacatca aagctgatatt cgctacgaaac 720 aaagataata ttgctaaaaaa agcaaacagt gccgacgtgt acaccagaga agagtctgac 780 agcaaaatttg tcagaatttga ttgtctgtcgtactaccg aaaaatttggg cacacgctt 840 gctctctgtc taaaatccat tgccgtatc gatactgcg tgaacggtt ggataaaaaca 900 gtgtcagacc tgcggcaaaa aacccggaa ggccttgcg aacaaggccg gctctccgg 960 ctgttccaaac ttacaacgtt gggtcgttca aatgtaacgg ctgcagtcgg cgctacaaaa 1020 tccgaatccgg cagtcggccat cggttccggc ttccgttta ccgaaaactt tgccgcca 1080 gcaggcgtgg cagtcggcact ttccgttccgg tccatccgt cgccgtcaat 1140 tacgagtggg gatccggccgg aggccggact tctgcgtcccg acttcaatgc aggccgttacc 1200 ggtatccggc gcaacacgac agcaacaaaaa gcaaaatcg cagcgtatc ttacgctt 1260 35 atcaagaacg aatgttgcgg agacagaaggc atgtctgtt ccgttccgg tgcgttgc 1320 gttacagaca gggatgccaat aatcaatgc ccccccccaat cttccatccat ccggagacttt 1380 ccaaacccaa atgacgcata caagaatttgc atcaacccat aacctcaat tgaagcaggc 1440 tatacaggac gccccgtttaga ggttaggtatc gtcgacacag gcaatccgt cgccagcata 1500 tcctttcccg aactgtatgg cagaaaaaaa caccgtata acgaaaatttcaaaaaacttat 1560 acggcgtata tgcggagggc agccgttgc gacggaggcc gtaaagacat tgaagctt 1620 40 ttcgacgtatg agggccgttat agagactgaa gcaaaaggccg cggatattccg ccacgtaaaa 1680 gaaatccggac acatcgattt ggtctccat attattggcc ggcgttccgt ggacggcaga 1740 cctgcaggcg gtattgcgcg cgatgcgacg ctacacataa tgaatacgaa tgatgaaacc 1800 aagaacgaaa ttagtggttgc agccatccgc aatgcgttgc tcaagctggg cgaacgtggc 1860 gtgcgcatcg tcaataacag ttttggaaaca acatcgaggc caggcaactgc cgaccctttc 1920 caaatacgcca attcggggc gcaatccgtc caacgttgc tcgactattt ccggcgttgc 1980 45 aaaacacgac aggtatccg cctgtatggc caagacgttgc acggcaactt gtcctaccac 2040 atccgtatataaaaatcgat tttcatctt tcgacaggca atgacgcaca agtcagccc 2100 aacacatatg ccctattggc attttatgaa aaagacgttc aaaaaggcat tattacagtc 2160 gcaggcgttag accgcgttgc agaaaaatgc aaacggggaa tgtatggaga accgggtaca 2220 50 gaaccgttgc agtatggctc caaccattgc ggaattactgc ccatgttgcgtt ccgttccgg 2280 ccctatggaa caagcgtccg tttccatccgt acaaaccggc ttcacatttc cggaaacatcc 2340 tttccgcac ccatcgtaac cggcaggccg gctctgtgc tgcaaaaaa cccgtggatg 2400 agcaacgaca accgtcgatc cacgttgcgt acgacggctc aggacatcg tgcgttgc 2460 55 gtggacagca agttccggctg gggactgtgc gatgcgggtt aggccatgaa cggacccggc 2520 tcctttccgt tcggcgactt taccgttgc acgaaaggta catccgtat tgcctactcc 2580 ttccgtatggc acatccgttgc caccggccg ctgatcaaaa aaggccggcag ccaactgca 2640 ctgcacggca acaacacacta tacggggcaaa accattatcg aaggccgttc gctgggttt 2700 tacggcaaca acaaatacgga tatgcgttgc gaaacccaaag gtgcgttgc ttataacggg 2760 gcggcatccg gccggcggctt gaaacggcgc ggcattgtct atctggcaga taccgaccaa 2820 tccggcgcac acgaaaccgtt acacatcaaa ggcgttgc acgtggacgg caaaggatcg 2880 ctgtacacac gtttggccaa actgttgcggta cggcgattat cggcggcaag 2940			

	ctgtacatgt	cggcacgcgg	caagggggca	ggcttatctca	acagtaccgg	acgacgtgtt	3000
	cccttcctga	gtgccgccaa	aatcgccag	gattattctt	tcttcacaaa	catcgaaacc	3060
5	gacggccgccc	tgctggcttc	cctcgacagc	gtcgaaaaaa	cagcgggcag	tgaaggcgcac	3120
	acgctgtcct	attatgtccg	tcgcggcaat	gcccgcacgg	ctgcttcggc	agccgcacat	3180
	tccgcggccc	ccggctctgaa	acacgcgtta	gaacaggcgc	cgacaaatct	ggaaaacctg	3240
	atggtcgaac	tggatgcctc	cgaatcatcc	gcaacaccccg	agacggttga	aactgcggca	3300
	gccgaccgca	cagatatgcc	gggcattccgc	ccctacggcg	caactttccg	cgcagcggca	3360
	gccgtacagc	atgcgaatgc	cgcccgcgtt	gtacgcacatc	tcaacagtct	cgccgcgtacc	3420
10	gtctatgccc	acagtaccgc	cgcccatgca	gatatgcagg	gacgcccct	gaaagccgta	3480
	tcggacgggt	tggaccacaa	cggcacgggt	ctgcgcgtca	tcgcgcacaa	ccaaacaggac	3540
	ggtggaaacgt	gggaaacaggg	cgtgttga	ggcaaaatcg	gccccgtac	ccaaaccgtc	3600
	ggcatttccg	cggaaaacccg	cgaaaatacg	acagcggccg	ccacacttgg	catgggacgc	3660
	agcacatgga	gcgaaaaacag	tgcaaatgca	aaaaccgaca	gcatttagtct	gtttcaggc	3720
	atacggcacc	atgcggcga	tatcggttat	ctcaaaaggcc	tgttctctta	cgacgcgtac	3780
15	aaaaacagca	tcagccgcag	caccggtgcg	gacgaacatg	cggaaggcag	cgtcaacggc	3840
	acgctgtatgc	agctggggcgc	actgggcgtt	gtcaacgttc	cgtttgcgc	aacgggagat	3900
	ttgacgtcg	aaggcggct	gctgtacgac	ctgtctaaac	aggatgcatt	cgccgaaaaaa	3960
	ggcagtcgtt	tgggctggag	cggtcaacagc	ctcaactgg	gcacgctgtt	cggactcgcg	4020
	ggtctgaagc	tgtcgcaacc	cttgcgtat	aaagccgtcc	tgttgcac	ggccggcgtg	4080
	gaacgcgacc	tgaacggacg	cgactacacg	gtaacggcgc	gctttaccgg	cgcgactgca	4140
	gcaacccgca	agacgggggc	acgcaatatg	ccgcacaccc	gtctggttgc	cgccctgggc	4200
20	gcccgtatcg	aattcggcaa	cggtctggaa	ggcttggcac	gttacagcta	cgccggttcc	4260
	aaacagtagc	gcaaccacag	cggtcgatc	ggcgttaggt	accggttct	cgagcaccac	4320
	caccaccacc	actgta					4335
	<210>	151					
	<211>	1444					
	<212>	PRT					
25	<213>	Artificial Sequence					
	<220>						
	<223>	961-983					
	<400>	151					
30	Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile						
	1	5	10	15			
	Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly						
	20	25	30				
	Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp						
35	35	40	45				
	Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu						
	50	55	60				
	Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln						
40	65	70	75	80			
	Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu						
	85	90	95				
	Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala						
45	100	105	110				
	Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile						
	115	120	125				
	Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu						
50	130	135	140				
	Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe						
	145	150	155	160			
	Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu						
	165	170	175				
55	Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys						
	180	185	190				

Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys  
 195 200 205  
 5 Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu  
 210 215 220  
 Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn  
 225 230 235 240  
 10 Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg  
 245 250 255  
 Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr  
 260 265 270  
 15 Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala  
 275 280 285  
 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu  
 290 295 300  
 20 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly  
 305 310 315 320  
 Leu Phe Gln Pro Tyr Asn Val Gly Arg Phe Asn Val Thr Ala Ala Val  
 325 330 335  
 25 Gly Gly Tyr Lys Ser Glu Ser Ala Val Ala Ile Gly Thr Gly Phe Arg  
 340 345 350  
 Phe Thr Glu Asn Phe Ala Ala Lys Ala Gly Val Ala Val Gly Thr Ser  
 355 360 365  
 30 Ser Gly Ser Ser Ala Ala Tyr His Val Gly Val Asn Tyr Glu Trp Gly  
 370 375 380  
 Ser Gly Gly Gly Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr  
 385 390 395 400  
 35 Gly Ile Gly Ser Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val  
 405 410 415  
 Ser Tyr Ala Gly Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu  
 420 425 430  
 40 Cys Ala Gly Arg Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile  
 435 440 445  
 Asn Ala Pro Pro Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn  
 450 455 460  
 Asp Ala Tyr Lys Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly  
 465 470 475 480  
 45 Tyr Thr Gly Arg Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser  
 485 490 495  
 Val Gly Ser Ile Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly  
 500 505 510  
 50 Tyr Asn Glu Asn Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala  
 515 520 525  
 Pro Glu Asp Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu  
 530 535 540  
 55 Ala Val Ile Glu Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys  
 545 550 555 560

Glu Ile Gly His Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser  
 565 570 575  
 5 Val Asp Gly Arg Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His  
 580 585 590  
 Ile Met Asn Thr Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala  
 595 600 605  
 10 Ile Arg Asn Ala Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val  
 610 615 620  
 Asn Asn Ser Phe Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe  
 625 630 635 640  
 15 Gln Ile Ala Asn Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr  
 645 650 655  
 Ser Gly Gly Asp Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser  
 660 665 670  
 20 Asp Tyr Gly Asn Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe  
 675 680 685  
 Ile Phe Ser Thr Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala  
 690 695 700  
 25 Leu Leu Pro Phe Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val  
 705 710 715 720  
 Ala Gly Val Asp Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly  
 725 730 735  
 30 Glu Pro Gly Thr Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile  
 740 745 750  
 Thr Ala Met Trp Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe  
 755 760 765  
 35 Thr Arg Thr Asn Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro  
 770 775 780  
 Ile Val Thr Gly Thr Ala Ala Leu Leu Leu Gln Lys Tyr Pro Trp Met  
 785 790 795 800  
 40 Ser Asn Asp Asn Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile  
 805 810 815  
 Gly Ala Val Gly Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala  
 820 825 830  
 45 Gly Lys Ala Met Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr  
 835 840 845  
 Ala Asp Thr Lys Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp  
 850 855 860  
 50 Ile Ser Gly Thr Gly Gly Leu Ile Lys Lys Gly Gly Ser Gln Leu Gln  
 865 870 875 880  
 Leu His Gly Asn Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Gly  
 885 890 895  
 55 Ser Leu Val Leu Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr  
 900 905 910  
 Lys Gly Ala Leu Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn  
 915 920 925

Ser Asp Gly Ile Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn  
 930 935 940  
 5 Glu Thr Val His Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr  
 945 950 955 960  
 Leu Tyr Thr Arg Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile  
 965 970 975  
 10 Ile Gly Gly Lys Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr  
 980 985 990  
 Leu Asn Ser Thr Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile  
 995 1000 1005  
 15 Gly Gln Asp Tyr Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu  
 1010 1015 1020  
 Leu Ala Ser Leu Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp  
 1025 1030 1035 1040  
 20 Thr Leu Ser Tyr Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser  
 1045 1050 1055  
 Ala Ala Ala His Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln  
 1060 1065 1070  
 25 Gly Gly Ser Asn Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu  
 1075 1080 1085  
 Ser Ser Ala Thr Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr  
 1090 1095 1100  
 30 Asp Met Pro Gly Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala  
 1105 1110 1115 1120  
 Ala Val Gln His Ala Asn Ala Ala Asp Gly Val Arg Ile Phe Asn Ser  
 1125 1130 1135  
 35 Leu Ala Ala Thr Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met  
 1140 1145 1150  
 Gln Gly Arg Arg Leu Lys Ala val Ser Asp Gly Leu Asp His Asn Gly  
 1155 1160 1165  
 40 Thr Gly Leu Arg Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp  
 1170 1175 1180  
 Glu Gln Gly Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val  
 1185 1190 1195 1200  
 45 Gly Ile Ala Ala Lys Thr Gly Glu Asn Thr Thr Ala Ala Ala Thr Leu  
 1205 1210 1215  
 Gly Met Gly Arg Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr  
 1220 1225 1230  
 50 Asp Ser Ile Ser Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile  
 1235 1240 1245  
 Gly Tyr Leu Lys Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile  
 1250 1255 1260  
 55 Ser Arg Ser Thr Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly  
 1265 1270 1275 1280  
 Thr Leu Met Gln Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala  
 1285 1290 1295

Ala Thr Gly Asp Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu  
 1300 1305 1310  
 5 Lys Gln Asp Ala Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly  
 1315 1320 1325  
 Asn Ser Leu Thr Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu  
 1330 1335 1340  
 10 Ser Gln Pro Leu Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val  
 1345 1350 1355 1360  
 Glu Arg Asp Leu Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr  
 1365 1370 1375  
 15 Gly Ala Thr Ala Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His  
 1380 1385 1390  
 Thr Arg Leu Val Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly  
 1395 1400 1405  
 20 Trp Asn Gly Leu Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly  
 1410 1415 1420  
 Asn His Ser Gly Arg Val Gly Val Gly Tyr Arg Phe Leu Glu His His  
 1425 1430 1435 1440  
 His His His His

25 <210> 152  
 <211> 2256  
 <212> DNA  
 <213> Artificial Sequence

30 <220>  
 <223> 961c-ORF46.1

35 <400> 152

atggccacaa	acgacgacga	tgttaaaaaa	gctgccactg	tggccattgc	tgctgcctac	60
aacaatggcc	aagaaaatcaa	cggtttcaaa	gctggagaga	ccatctacga	catttatgaa	120
gacggcacaa	ttaccaaaaa	agacgcaact	gcagccatg	ttgaagccga	cgactttaaa	180
ggtctgggtc	tgaaaaaaagt	cgtgactaac	ctgaccaaaa	ccgtcaatga	aaacaaacaa	240
aacgtcgatg	ccaaaggtaaa	agctgcagaa	tctgaaatag	aaaagttAAC	aaccaagttA	300
gcagacactg	atggcgctt	agcagatact	gatgccgtc	tggatgcAAC	caccaacGCC	360
ttgaataaat	tggagaaaaa	tataacgac	tttgcgtaa	agactaaAGAC	aatatcgTA	420
aaaattgtatg	aaaaatttaga	agccgtggt	gataccgtcg	acaaggatgc	cgaaggatTC	480
aacgatatcg	ccgattcatt	ggatgaaACC	aacactaagg	cagacgaAGC	cgtcaaaACC	540
gccaatgaag	ccaaacagac	ggccgaagaa	accaaacaaa	acgtcgatGC	caaagtAAAA	600
gctgcagaaa	ctgcagcagg	caaaggccaa	gctgccgtG	gcacagCTAA	tactgcAGCC	660
gacaaggccg	aagctgtcgc	tgcAAAAGT	accgacatCA	aagctgtAT	cgctacGAAC	720
aaagataata	ttgtctaaaa	agcaaACGT	gccgacgtgt	acaccAGAGA	agagtCTGAC	780
agcaaatttg	tcagaattga	tggctgtaa	gctactaccG	aaaaatttGGA	cacacGTTG	840
gcttctgtg	aaaaatccat	tgccgatCAC	gataactCGCC	tgaacGGTT	ggataaaaACA	900
gtgtcagacc	tgcCAAAGA	aacCCGCCAA	ggcTTGCA	aacaagCCGC	gcttCCGGT	960
ctgttccaa	cttacAAACGT	gggtggatCC	ggaggaggAG	gatcagATT	ggcaaACGAT	1020
tcttttatcc	ggcagggttC	cgaccgtcG	cattcgaAC	ccgacGGGAA	ataccaccta	1080
ttcggcagca	ggggggaaACT	tgcccggcgc	agcggccata	tcggattttGGG	aaaaataacaa	1140
agccatcagt	tggcaacCT	gatgattCA	caggcggCCA	ttaaaggAAA	tatcggCTAC	1200
attgtccgt	ttttcgatca	cgggcacGAA	gtccattCCC	ccttcgacAA	ccatgcCTCA	1260
cattccgatt	ctgatgaAGC	cggttagtCC	gttgacggat	ttagcTTA	ccgcacatCCAT	1320
tgggacggat	acgaacacca	tccccGCCAC	ggctatgacG	ggccacAGGG	cggcggCTAT	1380
cccgctccca	aaggcgcgag	ggatataAC	agctacgaca	taaaaggCGT	tgccccAAAT	1440
atccgcctca	acctgaccGA	caacccgcgc	accggacaAC	ggcttgcGA	ccgttCCAC	1500
aatggcggtA	gtatgtcGAC	gcaaggAGTA	ggcggacggat	tcaaaACGCGC	cacccgataAC	1560
agccccgagc	tggacagatC	gggcaatGCC	ggcgaAGGCT	tcaacGGCAC	tgcagatATC	1620
gttaaaaaaca	tcatcgGCgc	ggcaggAGAA	attgtcGGCG	caggcgtGTC	cgtgcaggGC	1680
ataagcgaag	gctcaaACAT	tgctgtCATG	cacggcttGG	gtctgCTTC	caccgaaaAC	1740

aagatggcgc gcatcaacga tttggcagat atggcgcaac tcaaagacta tgccgcagca 1800  
 gccatccgcg attgggcagt cccaaacccc aatgccgcac aaggcataga agccgtcagc 1860  
 5 aatatcttta tggcagccat ccccatcaa gggattggag ctgttcgggg aaaatacggc 1920  
 ttggggcggca tcacggcaca tcctatccaag cggtcgaga tgggcgcgat cgcatcgcc 1980  
 aaaggaaaaat ccgcgtcag cgacaatttt gccgatgcgg catacgccaa ataccgcgtcc 2040  
 ctttaccatt cccgaaatat cgttcaaac ttggagcgc gttacggcaa agaaaaacatc 2100  
 acctcctcaa ccgtgcccgc gtcaaacggc aaaaatgtca aactggcaga ccaacgccc 2160  
 ccgaagacag gcgtaccgtt tgacggtaaa gggtttccga attttgagaa gcacgtgaaa 2220  
 tatgatacgc tcgagcacca ccaccaccac cactga 2256

10 <210> 153  
 <211> 751  
 <212> PRT  
 <213> Artificial Sequence

15 <220>  
 <223> 961c-ORF46.1

20 <400> 153  
 Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile 15  
 1 Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly 20  
 25 Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp 30  
 35 35 40 45  
 25 Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu 50  
 55 60  
 65 Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln 70 75 80  
 30 85 90 95  
 35 Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu 100 105 110  
 115 120 125  
 40 Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala 130 135 140  
 145 150 155 160  
 45 Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu 165 170 175  
 180 185 190  
 50 Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys 195 200 205  
 210 215 220  
 55 Ala Val Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu 225 230 235 240  
 245 250 255  
 260 265 270

## EP 1 790 660 A2

Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala  
 275 280 285  
 5 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu  
 290 295 300  
 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly  
 305 310 315 320  
 10 Leu Phe Gln Pro Tyr Asn Val Gly Ser Gly Gly Ser Asp  
 325 330 335  
 Leu Ala Asn Asp Ser Phe Ile Arg Gln Val Leu Asp Arg Gln His Phe  
 340 345 350  
 15 Glu Pro Asp Gly Lys Tyr His Leu Phe Gly Ser Arg Gly Glu Leu Ala  
 355 360 365  
 Glu Arg Ser Gly His Ile Gly Leu Gly Lys Ile Gln Ser His Gln Leu  
 370 375 380  
 20 Gly Asn Leu Met Ile Gln Gln Ala Ala Ile Lys Gly Asn Ile Gly Tyr  
 385 390 395 400  
 Ile Val Arg Phe Ser Asp His Gly His Glu Val His Ser Pro Phe Asp  
 405 410 415  
 25 Asn His Ala Ser His Ser Asp Ser Asp Glu Ala Gly Ser Pro Val Asp  
 420 425 430  
 Gly Phe Ser Leu Tyr Arg Ile His Trp Asp Gly Tyr Glu His His Pro  
 435 440 445  
 30 Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly Tyr Pro Ala Pro Lys  
 450 455 460  
 Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile Lys Gly Val Ala Gln Asn  
 465 470 475 480  
 35 Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser Thr Gly Gln Arg Leu Ala  
 485 490 495  
 Asp Arg Phe His Asn Ala Gly Ser Met Leu Thr Gln Gly Val Gly Asp  
 500 505 510  
 40 Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro Glu Leu Asp Arg Ser Gly  
 515 520 525  
 Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala Asp Ile Val Lys Asn Ile  
 530 535 540  
 Ile Gly Ala Ala Gly Glu Ile Val Gly Ala Gly Asp Ala Val Gln Gly  
 545 550 555 560  
 45 Ile Ser Glu Gly Ser Asn Ile Ala Val Met His Gly Leu Gly Leu Leu  
 565 570 575  
 Ser Thr Glu Asn Lys Met Ala Arg Ile Asn Asp Leu Ala Asp Met Ala  
 580 585 590  
 50 Gln Leu Lys Asp Tyr Ala Ala Ala Ile Arg Asp Trp Ala Val Gln  
 595 600 605  
 Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala Val Ser Asn Ile Phe Met  
 610 615 620  
 55 Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala Val Arg Gly Lys Tyr Gly  
 625 630 635 640

Leu Gly Gly Ile Thr Ala His Pro Ile Lys Arg Ser Gln Met Gly Ala  
 645 650 655  
 5 Ile Ala Leu Pro Lys Gly Lys Ser Ala Val Ser Asp Asn Phe Ala Asp  
 660 665 670  
 Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr His Ser Arg Asn Ile Arg  
 675 680 685  
 10 Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu Asn Ile Thr Ser Ser Thr  
 690 695 700  
 Val Pro Pro Ser Asn Gly Lys Asn Val Lys Leu Ala Asp Gln Arg His  
 705 710 715 720  
 15 Pro Lys Thr Gly Val Pro Phe Asp Gly Lys Gly Phe Pro Asn Phe Glu  
 725 730 735  
 Lys His Val Lys Tyr Asp Thr Leu Glu His His His His His His  
 740 745 750  
 20 <210> 154  
 <211> 1773  
 <212> DNA  
 <213> Artificial Sequence  
 25 <220>  
 <223> 961c-741  
 <400> 154  
 atggccacaa acgacgacga tggtaaaaaaa gctgccactg tggccattgc tgctgcctac 60  
 aacaatggcc aagaatcaa cggtttcaaa gctggagaga ccatctacga cattgatgaa 120  
 gagccgcacaa ttaccaaaaa agacgcaact gcagccgatg ttgaagccga cgactttaaa 180  
 ggtctgggtc tgaaaaaaagt cgtgactaac ctgaccaaaa ccgtcaatga aaacaaacaa 240  
 aacgtcgatg ccaaagtaaa agctgcagaa tctgaaatag aaaagttAAC aaccaagttA 300  
 gcagacactg atggcgttt agcagatact gatgccgctc tggatgcaac caccAACGCC 360  
 ttgaataaat tgggagaaaa tataacgaca tttgctgaag agactaAGAC aaatATCGTA 420  
 aaaattgtatg aaaaatttaga agccgtgctg gataccgtcg acaaggatgc cgaagcatTC 480  
 aacgatatcg ccgatttcg ggtatggaaacc aacactaagg cagacgaaGC cgtccAAACC 540  
 gccaatgaag ccaaACAGAC ggCCGAAGAA accaaACAAA acgtcgatgc caaAGTAAAAA 600  
 gctgcagaaaa ctgcagcagg caaAGCCGAA gctgccgCTG gcacagctAA tactcAGCC 660  
 35 gacaaggccc aagctgtcgc tgcaaaAGTT accgacatCA aagctgatAT cgctacgAAC 720  
 aaagataata ttgtcaaaaa agcaaacAGT gccgacgtGT acaccAGAGA agagtCTGAC 780  
 agcaaaATTG tcagaATTG tggtctGAAC gctactACCG AAAAATTGGA cacACGCTTG 840  
 gcttctgtCTG aaaaATCCAT tgccgatCAC gataCTCGC tgaacCGTTT ggataAAACA 900  
 gtgtcagacc tgccaaAGA aACCCGCCAA ggcCTTGCAG aacaAGCCGc gcttCCGGT 960  
 40 ctgttccaaC cttacaacGT gggTggatCC ggAGGGGGTG gtgtcGCCGc cgacatCGGT 1020  
 gcggggcttg ccgatgcact aaccgcacCG ctcgaccata aagacaAAAGG tttcAGTCT 1080  
 ttgacgcTGG atcagtccgt caggaaaaAC gagAAAactGA agctggCGGC acaaggTGC 1140  
 gaaaaAACTT atggAAACGG tgacagCCTC aatacGGGCA aattGAAGAA cgacaaggTC 1200  
 agccgttCG actttatCCG ccaaaatCGAA gtggacGGGC agctcattAC tttggagAGT 1260  
 ggagagtTCC aagtatacaa acaaAGGcAT tccgcTTAA ccgcCTTCA gaccgAGCAA 1320  
 atacaaggat cggagcATT atagcggcgc aacatacATC ttttGACAAG cttcccGAAG gcccAGGGc gacatATCGC 1440  
 45 gggacggcgt tcggTTcaga cGATGCCGc ggAAAactGA CCTACACCAT agattcGCC 1500  
 gccaagcagg gaaacggcaa aatcgaACAT ttgaaATCGC cagaACTCAA tgtcGACCTG 1560  
 gcccGCCGc atatcaagCC ggtatggAAA CGCCATGCC tcatcAGCGG ttccGTCCTT 1620  
 tacaaccAAg ccgagAAAGG cagttaCTC CTCGGTATCT ttggcggAAA agcccAGGAA 1680  
 50 gttggcggca gCGCggAAgt gaaaACCGTA aacggcatac gccatATCGG ctttgccGCC 1740  
 aagcaactcg agcaccacca ccaccaccaC tga 1773  
 <210> 155  
 <211> 590  
 <212> PRT  
 <213> Artificial Sequence  
 55 <220>  
 <223> 961c-741

<400> 155  
 Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile  
 1 5 10 15  
 Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly  
 5 20 25 30  
 Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp  
 10 35 40 45  
 Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu  
 50 55 60  
 Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln  
 65 70 75 80  
 Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu  
 85 90 95  
 Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala  
 100 105 110  
 Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile  
 115 120 125  
 Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu  
 130 135 140  
 Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe  
 145 150 155 160  
 Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu  
 165 170 175  
 Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys  
 180 185 190  
 Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys  
 195 200 205  
 Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu  
 210 215 220  
 Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn  
 225 230 235 240  
 Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg  
 245 250 255  
 Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr  
 260 265 270  
 Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala  
 275 280 285  
 Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu  
 290 295 300  
 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly  
 305 310 315 320  
 Leu Phe Gln Pro Tyr Asn Val Gly Ser Gly Gly Gly Val Ala  
 325 330 335  
 Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala Leu Thr Ala Pro Leu Asp  
 340 345 350  
 His Lys Asp Lys Gly Leu Gln Ser Leu Thr Leu Asp Gln Ser Val Arg

	355	360	365	
5	Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln Gly Ala Glu Lys Thr Tyr			
	370	375	380	
	Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys Leu Lys Asn Asp Lys Val			
	385	390	395	400
10	Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu Val Asp Gly Gln Leu Ile			
	405	410	415	
	Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr Lys Gln Ser His Ser Ala			
	420	425	430	
15	Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln Asp Ser Glu His Ser Gly			
	435	440	445	
	Lys Met Val Ala Lys Arg Gln Phe Arg Ile Gly Asp Ile Ala Gly Glu			
	450	455	460	
20	His Thr Ser Phe Asp Lys Leu Pro Glu Gly Gly Arg Ala Thr Tyr Arg			
	465	470	475	480
	Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly Gly Lys Leu Thr Tyr Thr			
	485	490	495	
	Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly Lys Ile Glu His Leu Lys			
	500	505	510	
25	Ser Pro Glu Leu Asn Val Asp Leu Ala Ala Ala Asp Ile Lys Pro Asp			
	515	520	525	
	Gly Lys Arg His Ala Val Ile Ser Gly Ser Val Leu Tyr Asn Gln Ala			
	530	535	540	
30	Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe Gly Gly Lys Ala Gln Glu			
	545	550	555	560
	val Ala Gly Ser Ala Glu Val Lys Thr Val Asn Gly Ile Arg His Ile			
	565	570	575	
35	Gly Leu Ala Ala Lys Gln Leu Glu His His His His His His His			
	580	585	590	
	<210> 156			
	<211> 4170			
	<212> DNA			
40	<213> Artificial Sequence			
	<220>			
	<223> 961c-983			
	<400> 156			
45	atggccacaa acgacgacgaa tttaaaaaaa gctgccactg tggccattgc tgctgcctac			60
	aacaatggcc aagaaatcaa cggttcaaa gctggagaga ccatctacga cattgatgaa			120
	gacggcacaa ttacaaaaaa agacgcaact gcagccgatg ttgaagccga cgactttaaa			180
	ggtctgggtc taaaaaaaagt cgtgactaac ctgaccaaaa ccgtcaatga aaacaaacaa			240
	aacgtcgatg ccaaagtaaa agctgcagaa tctgaaatag aaaagttAAC aaccaagttA			300
	gcagacactg atgcccgttt agcagatact gatgccctc tggatgcaac cccaacGCC			360
	ttgaataaat tgggagaaaa tataacgaca ttgtctgaag agactaAGAC aaatatCGTA			420
50	aaaattgtatg aaaaatttaga agccgtggct gataccgtcg acaaggatgc cgaagcattc			480
	aacgatatacg ccgatttcatt ggatgaaacc aacactaagg cagacGAAGC cgtcaAAACC			540
	gccaatgaaAG CAAACACAGAC GGCCGAAAG ACCAAACAAA ACGTGATGC CAAAGTAAAAA			600
	gctgcagaaaa CTGCAAGCAGGG CAAAGCCGAA GCTGCCGCTG GCACAGCTAA TACTGCAGCC			660
	gacaaggccg aagctgtcgc tgcaaaagt accgacatca aagctgatAT CGCTACGAAAC			720
	aaagataata ttgctaaaaaa agcaaAGCTG GCGACGTGT ACACCAAGAGA AGAGTCTGAC			780
55	agcaaATTG tcagaattga ttgtctgaac gctactacGG AAAAATTGGA CACACGCTTG			840
	gcttcgtctg aaaaatccat tgccgatcac gatactcgcc tgaacggTTT ggataAAAACA			900
	gtgtcagacc tgcgcaaAGA AACCCGCCAA GGCCTTGAG aacaAGCCGC GCTCTCCGGT			960

	ctgttccaac	cttacaacgt	gggtggatcc	ggcggaggcg	gcacttctgc	gcccgaacttc	1020
5	aatgcagggcg	gtaccggtat	cggcagaac	agcagagcaa	caacagcga	atcagcagca	1080
	gtatcttacg	ccggtatcaa	gaacgaaatg	tgc当地agaca	gaagcatgtct	ctgtgccgt	1140
	cggatgcacg	ttgcggttac	agacaggat	gccaaatca	atgccccccc	cccgaatctg	1200
	cataccggag	acittccaaa	cccaaattac	gcatacaaga	atttgatcaa	cctcaaacct	1260
	gcaatttgaag	caggctatac	aggacgccc	gttagaggtag	gtatcgctga	cacaggcgaa	1320
	tccgtccgca	gcataatcctt	tcccgaaactg	tatggcagaa	aagaacacgg	ctataacgaa	1380
10	aattacaaa	actatacggc	gtatatgcgg	aaggaaagcgc	ctgaagacgg	aggcggtaaa	1440
	gacatttgaag	cttctttcga	cgataggccc	gttataagaga	ctgaagcaaa	gcccacggat	1500
	atccgcacag	taaaagaaaat	cgacacatc	gatttggct	cccatattat	tggcggggcgt	1560
	tccgtggacg	gcacgttcgc	aggcggtt	gcccgcgtg	cgacgtaca	cataatgaat	1620
	acgaatgtatg	aaaccaagaa	cgaaatgtatg	gttgcagcca	tccgcatac	atgggtcaag	1680
	ctggggcaac	gtggcgtgcg	catcgtaat	aacagtttt	gaacaacatc	gagggcaggc	1740
15	actgcccacc	ttttccaaat	agccaattcg	gaggagcagt	accgccaagc	gttgcgtcgc	1800
	tattccggcg	gtgataaaaac	agacgagggt	atccgcctga	tgcaacagag	cgattacggc	1860
	aacctgtcct	accacatccg	taataaaaaac	atgttttca	tctttcgac	aggcaatgac	1920
	gcacaacgtc	agcccaacac	atatgccta	ttggccatttt	atggaaaaga	cgctaaaaaaaa	1980
	ggcattatca	cagtgcgagg	cgtagaccgc	agtggagaaa	atgtccaaacg	gaaatgttat	2040
	ggagaaccgg	gtacagaacc	gcttggat	ggctccaaacc	attgcggat	tactgccatg	2100
20	tggtgcctgt	cggcacccct	tgaagcaagc	gtccgtttca	cccgtaaaaa	cccgattcaa	2160
	attgcccggaa	catcctttc	cgcacccatc	gtaaccggca	ccgcggctct	gctgctgcag	2220
	aaatacccg	ggtatgagcaa	cgacaacctg	cgtagaccgt	tgctgacgac	ggctcaggac	2280
	atcggtgcag	tcggcgttgc	cagaacttc	ggctggggac	tgctgtatgc	gggttaaggcc	2340
	atgaacggac	ccgcgttgc	ccgttgcgc	gactttaccg	ccgatacga	aggtacatcc	2400
	gatatttgcct	actccttccg	taacgacatt	tcaggcaccgg	gcccgttat	aaaaaaaggc	2460
	ggcagccaac	tgcactgtca	cggaacaaac	acctatacgg	gaaaaaccat	tatcgaaggc	2520
	ggttcgcctgg	tgttgtacgg	caacaacaaa	tcggatatgc	gcgtcgaac	caaagggtgc	2580
25	ctgatttata	acggggcggc	atccggcggc	agccctgaaca	gcgacggcat	tgttatctg	2640
	gcagataccg	accatccgg	cgcaaaacga	accgtacaca	tcaaggcag	tctgcagctg	2700
	gacggccaaag	gtacgtgt	cacatgttg	ggcaaaactgc	tgaaagggtg	cggtacggcg	2760
	attatccgcg	gcaagctgt	catgtccgca	cgccgcagg	gggcaggctt	tctcaacagt	2820
	accggacgac	gtgttccctt	cctgagtgcc	gccaaaatcg	ggcaggattt	ttctttcttc	2880
30	acaacacatcg	aaaccgacgg	cgccctgt	gcttccctcg	acagcgtcg	aaaaacagcg	2940
	ggcagtgaaag	gcgacacgc	gtccattat	gtccgtcgcg	gcaatgcggc	acggactgct	3000
	tcggcagcgg	cacattccgc	gcccgcgg	ctgaaaacacg	ccgtagaaca	gggcggcagc	3060
	aatctggaaa	acctgtatgt	cgaacttgat	gcctccgaaat	catccgcac	acccgagacg	3120
	gttggaaactg	cgccagccg	ccgcacat	atgcggggca	tccgccttca	cgccgcac	3180
35	ttccgcgcag	cgcgccgt	acagcatgc	aatggccccc	acgggtgtac	catcttcaac	3240
	agtctcgcgc	ctaccgtcta	tgccgacat	accggccccc	atgccatata	gcagggacgc	3300
	cgcctgaaag	ccgtatcg	cggttggac	cacaacggca	cggttctgc	cgtcatcg	3360
	caaaccac	aggacgggtgg	aacgtgggg	caggcgggtg	ttgaaggca	aatgcgcggc	3420
40	agtacccaaa	ccgtcgccat	tgccgcgaaa	accggcggaa	atacgacacg	agccgcacaca	3480
	ctgggcattgg	gacgcgcac	atggagcgg	aacagtgc	atgcacaaaac	cgacagcatt	3540
	agtctgttt	caggcatac	gcacgtgc	ggcgatata	gctatctca	aggcctgttc	3600
	tcctacggac	gtcacaaaa	cagcatc	cgacgcaccc	gtgcggacga	acatgcggaa	3660
	ggcagcgtca	acggcacgt	gatgcacgt	ggcgcactgg	cggtgtca	cgttccgttt	3720
	gcccgcacgg	gagatttgc	ggtcaaggc	ggtctgcgt	acgacctgt	caacacaggat	3780
	gcattcgcgc	aaaaaggcag	tgttttggc	tggagcggca	acagcctcac	tgaaggcag	3840
	ctggtcggac	tcgcgggtct	gaagctgtcg	caacccttga	gcgataaagc	cgtccgttt	3900
	gcaacggcgg	cgctggaaacg	cgacctgaa	ggacgcgact	acacggtaac	gggcggcttt	3960
45	accggcgcga	ctgcagcaac	cggaagacg	ggggcacgc	atatgcgc	caccgtctg	4020
	gttgcggcc	tgggcgcgg	tgtcaattc	ggcaacggct	ggaacggctt	ggcacgttac	4080
	agctacggcc	gttccaaaca	gtacggcaac	cacagcggac	gagtccggct	aggtacccgg	4140
	ttcctcgagc	accaccacca	ccaccactga				4170
	<210>	157					
	<211>	1389					
	<212>	PRT					
	<213>	Artificial Sequence					
50	<220>						
	<223>	961c-983					
	<400>	157					
	Met Ala Thr Asn Asp Asp Asp Val Lys Lys Ala Ala Thr Val Ala Ile						
	1	5	10	15			
55	Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile Asn Gly Phe Lys Ala Gly						
	20	25	30				

Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly Thr Ile Thr Lys Lys Asp  
 35 40 45

5 Ala Thr Ala Ala Asp Val Glu Ala Asp Asp Phe Lys Gly Leu Gly Leu  
 50 55 60

Lys Lys Val Val Thr Asn Leu Thr Lys Thr Val Asn Glu Asn Lys Gln  
 65 70 75 80

10 Asn Val Asp Ala Lys Val Lys Ala Ala Glu Ser Glu Ile Glu Lys Leu  
 85 90 95

Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala Leu Ala Asp Thr Asp Ala  
 100 105 110

15 Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn Lys Leu Gly Glu Asn Ile  
 115 120 125

Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn Ile Val Lys Ile Asp Glu  
 130 135 140

20 Lys Leu Glu Ala Val Ala Asp Thr Val Asp Lys His Ala Glu Ala Phe  
 145 150 155 160

Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr Asn Thr Lys Ala Asp Glu  
 165 170 175

25 Ala Val Lys Thr Ala Asn Glu Ala Lys Gln Thr Ala Glu Glu Thr Lys  
 180 185 190

Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu Thr Ala Ala Gly Lys  
 195 200 205

30 Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr Ala Ala Asp Lys Ala Glu  
 210 215 220

Ala Val Ala Ala Lys Val Thr Asp Ile Lys Ala Asp Ile Ala Thr Asn  
 225 230 235 240

35 Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser Ala Asp Val Tyr Thr Arg  
 245 250 255

Glu Glu Ser Asp Ser Lys Phe Val Arg Ile Asp Gly Leu Asn Ala Thr  
 260 265 270

40 Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser Ala Glu Lys Ser Ile Ala  
 275 280 285

Asp His Asp Thr Arg Leu Asn Gly Leu Asp Lys Thr Val Ser Asp Leu  
 290 295 300

45 Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu Gln Ala Ala Leu Ser Gly  
 305 310 315 320

Leu Phe Gln Pro Tyr Asn Val Gly Ser Gly Gly Gly Thr Ser  
 325 330 335

50 Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly Ile Gly Ser Asn Ser Arg  
 340 345 350

Ala Thr Thr Ala Lys Ser Ala Ala Val Ser Tyr Ala Gly Ile Lys Asn  
 355 360 365

Glu Met Cys Lys Asp Arg Ser Met Leu Cys Ala Gly Arg Asp Asp Val  
 370 375 380

55 Ala Val Thr Asp Arg Asp Ala Lys Ile Asn Ala Pro Pro Pro Asn Leu  
 385 390 395 400

His Thr Gly Asp Phe Pro Asn Pro Asn Asp Ala Tyr Lys Asn Leu Ile  
 405 410 415  
 5 Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr Thr Gly Arg Gly Val Glu  
 420 425 430  
 Val Gly Ile Val Asp Thr Gly Glu Ser Val Gly Ser Ile Ser Phe Pro  
 435 440 445  
 10 Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr Asn Glu Asn Tyr Lys Asn  
 450 455 460  
 Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro Glu Asp Gly Gly Lys  
 465 470 475 480  
 15 Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala Val Ile Glu Thr Glu Ala  
 485 490 495  
 Lys Pro Thr Asp Ile Arg His Val Lys Glu Ile Gly His Ile Asp Leu  
 500 505 510  
 20 Val Ser His Ile Ile Gly Gly Arg Ser Val Asp Gly Arg Pro Ala Gly  
 515 520 525  
 Gly Ile Ala Pro Asp Ala Thr Leu His Ile Met Asn Thr Asn Asp Glu  
 530 535 540  
 25 Thr Lys Asn Glu Met Met Val Ala Ala Ile Arg Asn Ala Trp Val Lys  
 545 550 555 560  
 Leu Gly Glu Arg Gly Val Arg Ile Val Asn Asn Ser Phe Gly Thr Thr  
 565 570 575  
 30 Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln Ile Ala Asn Ser Glu Glu  
 580 585 590  
 Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser Gly Gly Asp Lys Thr Asp  
 595 600 605  
 35 Glu Gly Ile Arg Leu Met Gln Gln Ser Asp Tyr Gly Asn Leu Ser Tyr  
 610 615 620  
 His Ile Arg Asn Lys Asn Met Leu Phe Ile Phe Ser Thr Gly Asn Asp  
 625 630 635 640  
 40 Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu Leu Pro Phe Tyr Glu Lys  
 645 650 655  
 Asp Ala Gln Lys Gly Ile Ile Thr Val Ala Gly Val Asp Arg Ser Gly  
 660 665 670  
 45 Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu Pro Gly Thr Glu Pro Leu  
 675 680 685  
 Glu Tyr Gly Ser Asn His Cys Gly Ile Thr Ala Met Trp Cys Leu Ser  
 690 695 700  
 50 Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr Arg Thr Asn Pro Ile Gln  
 705 710 715 720  
 Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile Val Thr Gly Thr Ala Ala  
 725 730 735  
 Leu Leu Leu Gln Lys Tyr Pro Trp Met Ser Asn Asp Asn Leu Arg Thr  
 740 745 750  
 55 Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly Ala Val Gly Val Asp Ser  
 755 760 765

## EP 1 790 660 A2

Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly Lys Ala Met Asn Gly Pro  
 770 775 780  
 5 Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala Asp Thr Lys Gly Thr Ser  
 785 790 795 800  
 Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile Ser Gly Thr Gly Leu  
 805 810 815  
 10 Ile Lys Lys Gly Gly Ser Gln Leu Gln Leu His Gly Asn Asn Thr Tyr  
 820 825 830  
 Thr Gly Lys Thr Ile Ile Glu Gly Gly Ser Leu Val Leu Tyr Gly Asn  
 835 840 845  
 15 Asn Lys Ser Asp Met Arg Val Glu Thr Lys Gly Ala Leu Ile Tyr Asn  
 850 855 860  
 Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser Asp Gly Ile Val Tyr Leu  
 865 870 875 880  
 20 Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu Thr Val His Ile Lys Gly  
 885 890 895  
 Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu Tyr Thr Arg Leu Gly Lys  
 900 905 910  
 25 Leu Leu Lys Val Asp Gly Thr Ala Ile Ile Gly Gly Lys Leu Tyr Met  
 915 920 925  
 Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu Asn Ser Thr Gly Arg Arg  
 930 935 940  
 30 Val Pro Phe Leu Ser Ala Ala Lys Ile Gly Gln Asp Tyr Ser Phe Phe  
 945 950 955 960  
 Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu Ala Ser Leu Asp Ser Val  
 965 970 975  
 35 Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr Leu Ser Tyr Tyr Val Arg  
 980 985 990  
 Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala Ala His Ser Ala Pro  
 995 1000 1005  
 40 Ala Gly Leu Lys His Ala Val Glu Gln Gly Gly Ser Asn Leu Glu Asn  
 1010 1015 1020  
 Leu Met Val Glu Leu Asp Ala Ser Glu Ser Ser Ala Thr Pro Glu Thr  
 1025 1030 1035 1040  
 45 Val Glu Thr Ala Ala Ala Asp Arg Thr Asp Met Pro Gly Ile Arg Pro  
 1045 1050 1055  
 Tyr Gly Ala Thr Phe Arg Ala Ala Ala Ala val Gln His Ala Asn Ala  
 1060 1065 1070  
 Ala Asp Gly Val Arg Ile Phe Asn Ser Leu Ala Ala Thr Val Tyr Ala  
 1075 1080 1085  
 50 Asp Ser Thr Ala Ala His Ala Asp Met Gln Gly Arg Arg Leu Lys Ala  
 1090 1095 1100  
 Val Ser Asp Gly Leu Asp His Asn Gly Thr Gly Leu Arg Val Ile Ala  
 1105 1110 1115 1120  
 55 Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu Gln Gly Gly Val Glu Gly  
 1125 1130 1135

Lys Met Arg Gly Ser Thr Gln Thr Val Gly Ile Ala Ala Lys Thr Gly  
 1140 1145 1150  
 5 Glu Asn Thr Thr Ala Ala Ala Thr Leu Gly Met Gly Arg Ser Thr Trp  
 1155 1160 1165  
 Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp Ser Ile Ser Leu Phe Ala  
 1170 1175 1180  
 10 Gly Ile Arg His Asp Ala Gly Asp Ile Gly Tyr Leu Lys Gly Leu Phe  
 1185 1190 1195 1200  
 Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser Arg Ser Thr Gly Ala Asp  
 1205 1210 1215  
 15 Glu His Ala Glu Gly Ser Val Asn Gly Thr Leu Met Gln Leu Gly Ala  
 1220 1225 1230  
 Leu Gly Gly Val Asn Val Pro Phe Ala Ala Thr Gly Asp Leu Thr Val  
 1235 1240 1245  
 20 Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys Gln Asp Ala Phe Ala Glu  
 1250 1255 1260  
 Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn Ser Leu Thr Glu Gly Thr  
 1265 1270 1275 1280  
 25 Leu Val Gly Leu Ala Gly Leu Lys Leu Ser Gln Pro Leu Ser Asp Lys  
 1285 1290 1295  
 Ala Val Leu Phe Ala Thr Ala Gly Val Glu Arg Asp Leu Asn Gly Arg  
 1300 1305 1310  
 30 Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly Ala Thr Ala Ala Thr Gly  
 1315 1320 1325  
 Lys Thr Gly Ala Arg Asn Met Pro His Thr Arg Leu Val Ala Gly Leu  
 35 1330 1335 1340  
 Gly Ala Asp Val Glu Phe Gly Asn Gly Trp Asn Gly Leu Ala Arg Tyr  
 1345 1350 1355 1360  
 Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn His Ser Gly Arg Val Gly  
 1365 1370 1375  
 40 Val Gly Tyr Arg Phe Leu Glu His His His His His His  
 1380 1385  
 <210> 158  
 <211> 2304  
 <212> DNA  
 <213> Artificial Sequence  
 45 <220>  
 <223> 961cL-ORF46.1  
 <400> 158  
 atgaaacact ttccatccaa agtactgacc acagccatcc ttgccactt ctgttagccgc 60  
 gcactggcag ccacaaacga cgacgatgtt aaaaaagctg ccactgtggc cattgctgct 120  
 gcctacaaca atggccaaga aatcaacggt ttcaaagctg gagagaccat ctacgacatt 180  
 gatgaagacg gcacaattac caaaaaagac gcaactgcag ccgatgttg agccgacgac 240  
 tttaaaggtc tgggtctgaa aaaagtctgt actaacctga ccaaaaacgt caatgaaaac 300  
 aaacaaaacg tcgatgccaa agtaaaaagct gcagaatctg aaatagaaaa gttaacaacc 360  
 aagtttagcag acatcgatgc cgctttagca gatactgatg ccgctctgg tgcaaccacc 420  
 aacgccttga ataaatttggg agaaaaatata acgacattt ctgaagagac taagacaaaat 480  
 50 atcgtaaaaa ttgataaaa attagaagcc gtggctgata ccgtcgacaa gcatgccgaa 540  
 gcattcaacg atatcgccga ttcatggat gaaaccaaca ctaaggcaga cgaaggcgtc 600  
 aaaaccgcca atgaagccaa acagacggcc gaagaaacca aacaaaacgt cgatgccaaa 660

	gtaaaagctg cagaaactgc agcaggcaaa gccgaagctg ccgctggcac agctaatact	720
5	gcagccgaca aggccgaaagc tgcgcgtca aaagtttacgg acatcaaagc tgatatacgct	780
	acgaacaaag ataataattgc taaaaaaagca aacagtgcgg acgtgtacac cagagaagag	840
	tctgacagca aatttgcgt aattgtatgg ctgaacgcta ctaccgaaaa attggacaca	900
	cgcttgctt ctgctgaaaa atccattgcc gatcacgata tcgcctgaa cggtttggat	960
	aaaacagtgt cagacctgca caaagaaacc cgccaaggcc ttgcagaaca agccgcgctc	1020
	tccggctgt tccaacctta caacgtgggt ggatccggag gaggaggatc agatttggca	1080
10	aacattctt ttatccggca gggtctcgac cgtcagcatt tcgaacccga cggaaatac	1140
	cacctattcg gcagcagggg ggaacttgcc gagcgcagcg gccatatcg attggggaaa	1200
	atacaaagcc atcagtggg caacctgatg attcaacagg cggcattaa aggaaatac	1260
	ggctacattg tccgttttc cgatcacggg cacaaggctt attccccctt cgacaaccat	1320
	gcctcacatt ccgattctga tgaaggccgt agtcccgtt acggatttag ccttaccgc	1380
	atccatggg acggatacga acaccatccc gcccacggct atgacgggcc acagggcggc	1440
15	ggctatcccgg ctcccaaagg cgcgagggtt atatacagct acgacataaa aggcttgcc	1500
	caaaatatcc gcctcaacct gaccgacaac cgacgaccg gacaacgct tgccgaccgt	1560
	ttccacaatg ccgttagtat gctgacgcaa ggagttaggc acggattcaa acgcgcccacc	1620
	cgatacagcc ccgagctgga cagatcgggc aatgcccggc aagccttcaa cggcaactgca	1680
	gatatcgta aaaacatcat cggcgcggca ggagaaattt tcggcgcagg cgtgcccgt	1740
	caggcataa gcaaggctc aaacattgt gtcatgcacg gcttgggtct gctttccacc	1800
	gaaaacaaga tggcgcgcac caacgattt gcagatattgg cgcaactcaa agactatgcc	1860
	gcagcagcca tccgcgattt ggcagtccaa aaccccaatg cgcacacaagg catagaagcc	1920
20	gtcagcaata tctttatggc agccatcccc atcaaaggga ttggagctgt tcggggaaa	1980
	tacggcttgg gcccacatcac ggcacatctt atcaagcggt cgcagatgg cgcatcgca	2040
	ttgcccggaaag ggaatccgc cgtcagcgc aattttggc atgcggcata cgcacaaatac	2100
	ccgtccctt accattcccg aaatattcg tcaaacttgg agcagcgta cggcaaaagaa	2160
	aacatcacct cctcaaccgt gcccgcgtca aacggcaaaa atgtcaacta ggcagaccaa	2220
	cgccacccga agacaggcgat accgtttgac ggttaagggt ttccgaattt tgagaagcac	2280
	tgaaatatg atacgttaact cgag	2304
25	<210> 159	
	<211> 765	
	<212> PRT	
	<213> Artificial Sequence	
30	<220>	
	<223> 961cL-ORF46.1	
	<400> 159	
	Met Lys His Phe Pro Ser Lys Val Leu Thr Thr Ala Ile Leu Ala Thr	
	1 5 10 15	
35	Phe Cys Ser Gly Ala Leu Ala Ala Thr Asn Asp Asp Asp Val Lys Lys	
	20 25 30	
	Ala Ala Thr Val Ala Ile Ala Ala Tyr Asn Asn Gly Gln Glu Ile	
	35 40 45	
40	Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly	
	50 55 60	
	Thr Ile Thr Lys Lys Asp Ala Thr Ala Ala Asp Val Glu Ala Asp Asp	
	65 70 75 80	
45	Phe Lys Gly Leu Gly Leu Lys Val Val Thr Asn Leu Thr Lys Thr	
	85 90 95	
	Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu	
	100 105 110	
50	Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala	
	115 120 125	
	Leu Ala Asp Thr Asp Ala Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn	
	130 135 140	
55	Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn	
	145 150 155 160	
	Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala Asp Thr Val Asp	

	165	170	175
5	Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr 180 185 190		
	Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala Asn Glu Ala Lys Gln 195 200 205		
10	Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala 210 215 220		
	Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr 225 230 235 240		
15	Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys Val Thr Asp Ile Lys 245 250 255		
	Ala Asp Ile Ala Thr Asn Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser 260 265 270		
20	Ala Asp Val Tyr Thr Arg Glu Glu Ser Asp Ser Lys Phe Val Arg Ile 275 280 285		
	Asp Gly Leu Asn Ala Thr Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser 290 295 300		
25	Ala Glu Lys Ser Ile Ala Asp His Asp Thr Arg Leu Asn Gly Leu Asp 305 310 315 320		
	Lys Thr Val Ser Asp Leu Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu 325 330 335		
30	Gln Ala Ala Leu Ser Gly Leu Phe Gln Pro Tyr Asn Val Gly Gly Ser 340 345 350		
	Gly Gly Gly Ser Asp Leu Ala Asn Asp Ser Phe Ile Arg Gln Val 355 360 365		
35	Leu Asp Arg Gln His Phe Glu Pro Asp Gly Lys Tyr His Leu Phe Gly 370 375 380		
	Ser Arg Gly Glu Leu Ala Glu Arg Ser Gly His Ile Gly Leu Gly Lys 385 390 395 400		
40	Ile Gln Ser His Gln Leu Gly Asn Leu Met Ile Gln Gln Ala Ala Ile 405 410 415		
	Lys Gly Asn Ile Gly Tyr Ile Val Arg Phe Ser Asp His Gly His Glu 420 425 430		
45	Val His Ser Pro Phe Asp Asn His Ala Ser His Ser Asp Ser Asp Glu 435 440 445		
	Ala Gly Ser Pro Val Asp Gly Phe Ser Leu Tyr Arg Ile His Trp Asp 450 455 460		
50	Gly Tyr Glu His His Pro Ala Asp Gly Tyr Asp Gly Pro Gln Gly Gly 465 470 475 480		
	Gly Tyr Pro Ala Pro Lys Gly Ala Arg Asp Ile Tyr Ser Tyr Asp Ile 485 490 495		
55	Lys Gly Val Ala Gln Asn Ile Arg Leu Asn Leu Thr Asp Asn Arg Ser 500 505 510		
	Thr Gly Gln Arg Leu Ala Asp Arg Phe His Asn Ala Gly Ser Met Leu 515 520 525		
	Thr Gln Gly Val Gly Asp Gly Phe Lys Arg Ala Thr Arg Tyr Ser Pro		

	530	535	540														
5	Glu Leu Asp Arg Ser Gly Asn Ala Ala Glu Ala Phe Asn Gly Thr Ala 545 550 555 560																
	Asp Ile Val Lys Asn Ile Ile Gly Ala Ala Gly Glu Ile Val Gly Ala 565 570 575																
10	Gly Asp Ala Val Gln Gly Ile Ser Glu Gly Ser Asn Ile Ala Val Met 580 585 590																
	His Gly Leu Gly Leu Leu Ser Thr Glu Asn Lys Met Ala Arg Ile Asn 595 600 605																
15	Asp Leu Ala Asp Met Ala Gln Leu Lys Asp Tyr Ala Ala Ala Ile 610 615 620																
	Arg Asp Trp Ala Val Gln Asn Pro Asn Ala Ala Gln Gly Ile Glu Ala 625 630 635 640																
	Val Ser Asn Ile Phe Met Ala Ala Ile Pro Ile Lys Gly Ile Gly Ala 645 650 655																
20	Val Arg Gly Lys Tyr Gly Leu Gly Ile Thr Ala His Pro Ile Lys 660 665 670																
	Arg Ser Gln Met Gly Ala Ile Ala Leu Pro Lys Gly Lys Ser Ala Val 675 680 685																
25	Ser Asp Asn Phe Ala Asp Ala Ala Tyr Ala Lys Tyr Pro Ser Pro Tyr 690 695 700																
	His Ser Arg Asn Ile Arg Ser Asn Leu Glu Gln Arg Tyr Gly Lys Glu 705 710 715 720																
30	Asn Ile Thr Ser Ser Thr Val Pro Pro Ser Asn Gly Lys Asn Val Lys 725 730 735																
	Leu Ala Asp Gln Arg His Pro Lys Thr Gly Val Pro Phe Asp Gly Lys 740 745 750																
35	Gly Phe Pro Asn Phe Glu Lys His Val Lys Tyr Asp Thr 755 760 765																
	<210> 160																
	<211> 1839																
	<212> DNA																
40	<213> Artificial Sequence																
	<220>																
	<223> 961cL-741																
	<400> 160																
45	atgaaacact ttccatccaa agtactgacc acagccatcc ttgccacttt ctgtagccgc gcactggcag ccacaaacga cgacgatgtt aaaaaagctg ccactgtggc cattgctgct gcctacaaca atggccaaga aatcaacggt ttcaaagctg gagagaccat ctacgacatt gatgaagacg gcacaattac caaaaaagac gcaactgcag ccgatgttga agccgacgac tttaaaggctc tgggtctgaa aaaagtctgt actaacctga ccaaaaaccgt caatgaaaac aaacaaaacg tcgatgccaa agtaaaagctg gcagaatctg aaatagaaaa gttAACAAACC aagtttagcag acactgtatgc cgcttttagca gatactgtatgc ccgctctggaa tgcaaccacc aacgccttga ataaattggg agaaaaatata accacattt ctgaagagac taagacaaat atcgtaaaaa ttgatgaaaa attagaagcc gtggctgata ccgtcgacaa gcatgccgaa gcattcaacg atatcgccga ttcattggat gaaaccaaca ctaaggcaga cgaagccgtc aaaacccgcca atgaagccaa acagacggcc gaagaaacca aacaaaacgt cgatgccaaa gtaaaagctg cagaaactgc agcaggcaaa gcccgaagctg ccgctggcac agctaataact gcagccgaca aggccgaaagc tgcgtcgatca aagttaccg acatcaaaacg tgatatcgct acgaacaaag ataataattgc taaaaaaagca aacagtgcgc acgtgtacac cagagaagag tctgacagca aatttgtcag aattgtatgtt ctgaacgcata ctaccggaaa attggacaca cgcttggctt ctgctgaaaa atccattgcc gatcacgata ctcgcctgaa cggtttggat	60	120	180	240	300	360	420	480	540	600	660	720	780	840	900	960

5       aaaacagtgt cagacctgca caaagaaacc cgccaaggcc ttgcagaaca agccgcgtc   1020  
 tccggcttgt tccaaccttca acacgtgggt ggatccggag ggggtggtgt cgccgcgc   1080  
 atcggtgcgg ggcttgccga tgcaactaacc gcacccgtcg accataaaga caaaggttt   1140  
 cagtcttga cgctggatca gtccgtcagg aaaaacgaga aactgaagct ggcggcacaa   1200  
 ggtgcggaaa aaacttatgg aaacgggtac agcctcaata cgggcaatt gaagaacgac   1260  
 aaggtcagcc gtttcgactt tatccgcca atcgaagtgg acgggcagct cattacctt   1320  
 gagagtggag agttccaagt atacaaacaa agcattccg ctttaaccgc ctttcagacc   1380  
 gagcaaatac aagattcggc gcattccggg aagatgggt gcaaacgcgc gttcagaatc   1440  
 ggcgacatag cggcgaaca tacatcttt gacaagcttc ccgaaggccg cagggcgaca   1500  
 tattcgcggg cggcgttcgg ttcagacgt gcccgggaa aactgaccta caccatagat   1560  
 ttgcggcca agcaggggaaa cggcaaaatc gaacatttgaa atcgcgcaga actcaatgtc   1620  
 gacctggccg cggcgatat caagccggat ggaaaacgcgc atgcccgtcat cagcggttcc   1680  
 gtccttaca accaagccga gaaaggcagt tactccctcg gtatcttgg cgaaaaagcc   1740  
 caggaagttg cggcagcgc ggaagtggaa accgtaaacgc gcatacgcgc tatcggcctt   1800  
 gccgccaagc aactcgagca ccaccaccac caccactgaa                           1839

15      <210> 161  
 <211> 612  
 <212> PRT  
 <213> Artificial Sequence

20      <220>  
 <223> 961cL-741

25      <400> 161  
 Met Lys His Phe Pro Ser Lys Val Leu Thr Thr Ala Ile Leu Ala Thr  
 1                   5                   10                   15

30      Phe Cys Ser Gly Ala Leu Ala Ala Thr Asn Asp Asp Asp Val Lys Lys  
 20                   25                   30

35      Ala Ala Thr Val Ala Ile Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile  
 35                   40                   45

40      Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly  
 50                   55                   60

45      Thr Ile Thr Lys Lys Asp Ala Thr Ala Ala Asp Val Glu Ala Asp Asp  
 65                   70                   75                   80

50      Phe Lys Gly Leu Gly Leu Lys Lys Val Val Thr Asn Leu Thr Lys Thr  
 85                   90                   95

55      Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu  
 100                  105                  110

60      Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala  
 115                  120                  125

65      Leu Ala Asp Thr Asp Ala Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn  
 130                  135                  140

70      Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn  
 145                  150                  155                  160

75      Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala Asp Thr Val Asp  
 165                  170                  175

80      Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr  
 180                  185                  190

85      Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala Asn Glu Ala Lys Gln  
 195                  200                  205

90      Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala  
 210                  215                  220

95      Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Gly Thr Ala Asn Thr  
 225                  230                  235                  240

Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys Val Thr Asp Ile Lys  
 245 250 255  
 5 Ala Asp Ile Ala Thr Asn Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser  
 260 265 270  
 Ala Asp Val Tyr Thr Arg Glu Glu Ser Asp Ser Lys Phe Val Arg Ile  
 275 280 285  
 10 Asp Gly Leu Asn Ala Thr Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser  
 290 295 300  
 Ala Glu Lys Ser Ile Ala Asp His Asp Thr Arg Leu Asn Gly Leu Asp  
 305 310 315 320  
 15 Lys Thr Val Ser Asp Leu Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu  
 325 330 335  
 Gln Ala Ala Leu Ser Gly Leu Phe Gln Pro Tyr Asn Val Gly Gly Ser  
 340 345 350  
 20 Gly Gly Gly Gly Val Ala Ala Asp Ile Gly Ala Gly Leu Ala Asp Ala  
 355 360 365  
 Leu Thr Ala Pro Leu Asp His Lys Asp Lys Gly Leu Gln Ser Leu Thr  
 370 375 380  
 25 Leu Asp Gln Ser Val Arg Lys Asn Glu Lys Leu Lys Leu Ala Ala Gln  
 385 390 395 400  
 Gly Ala Glu Lys Thr Tyr Gly Asn Gly Asp Ser Leu Asn Thr Gly Lys  
 405 410 415  
 30 Leu Lys Asn Asp Lys Val Ser Arg Phe Asp Phe Ile Arg Gln Ile Glu  
 420 425 430  
 Val Asp Gly Gln Leu Ile Thr Leu Glu Ser Gly Glu Phe Gln Val Tyr  
 435 440 445  
 35 Lys Gln Ser His Ser Ala Leu Thr Ala Phe Gln Thr Glu Gln Ile Gln  
 450 455 460  
 Asp Ser Glu His Ser Gly Lys Met Val Ala Lys Arg Gln Phe Arg Ile  
 465 470 475 480  
 40 Gly Asp Ile Ala Gly Glu His Thr Ser Phe Asp Lys Leu Pro Glu Gly  
 485 490 495  
 Gly Arg Ala Thr Tyr Arg Gly Thr Ala Phe Gly Ser Asp Asp Ala Gly  
 500 505 510  
 45 Gly Lys Leu Thr Tyr Thr Ile Asp Phe Ala Ala Lys Gln Gly Asn Gly  
 515 520 525  
 Lys Ile Glu His Leu Lys Ser Pro Glu Leu Asn Val Asp Leu Ala Ala  
 530 535 540  
 50 Ala Asp Ile Lys Pro Asp Gly Lys Arg His Ala Val Ile Ser Gly Ser  
 545 550 555 560  
 Val Leu Tyr Asn Gln Ala Glu Lys Gly Ser Tyr Ser Leu Gly Ile Phe  
 565 570 575  
 Gly Gly Lys Ala Gln Glu Val Ala Gly Ser Ala Glu Val Lys Thr Val  
 580 585 590  
 55 Asn Gly Ile Arg His Ile Gly Leu Ala Ala Lys Gln Leu Glu His His  
 595 600 605

	His	His	His	His	610		
5	<210>	162					
	<211>	4218					
	<212>	DNA					
	<213>	Artificial Sequence					
10	<220>						
	<223>	961cL-983					
	<400>	162					
	atgaaacact	ttccatccaa	agtactgacc	acagccatcc	ttgccacttt	ctgtacggc	60
	gcactggcag	ccacaaacga	cgacgatgtt	aaaaaagctg	ccactgtggc	cattgctgtc	120
	gcctacaaca	atggccaaga	aatcaacggt	ttcaaagctg	gagagaccat	ctacgacatt	180
	gatgaagacg	gcacaattac	aaaaaaagac	gcaactgcag	ccgatgttga	agccgacgac	240
	tttaaaggtc	tgggtctgaa	aaaagtctgt	actaacctga	ccaaaaccgt	caatgaaaac	300
	aaacaaaacg	tcgatgcca	agtaaaagct	gcagaatctg	aaatagaaaa	gttaacaacc	360
	aagttagcag	actatgtc	cgctttagca	gatactgtat	ccgcttcttga	tgcacaccac	420
	aacgccttgc	ataaaattggg	agaaaaatata	acgacatttg	ctgaagagac	taagacaat	480
	atcgtaaaaa	ttgtatgaaaa	attagaagcc	ttgtgttata	ccgtcgacaa	gcatgcccac	540
	gcattcaacg	atatcgccg	ttcattttgtat	gaaaccaaca	ctaaggcaga	cgaagccgtc	600
	aaaaccgcca	atgaagccaa	acagacggcc	gaagaaacca	aacaaaacgt	cgatgcca	660
	gtaaaagctg	cagaaactgc	agcaggccaa	gccgaagctg	ccgcttgcac	agctaatact	720
	gcagccgaca	aggccgaaacg	tgtcgttgc	aaagtttaccg	acatcaaagc	tgatatcgct	780
	acgaacaaag	ataatattgc	aaaaaaagca	aacatgtccg	acgtgtacac	cagagaagag	840
	tctgacgacg	aatttgtcag	aattgtatgtt	ctgaaacgtt	ctaccggaaaa	attggacaca	900
	cgcttgcctt	ctgtatgaaaa	atccatttgc	gatcacgata	ctcgccctgaa	cggtttggat	960
	aaaacagtgt	cagacctgc	caaagaaacc	cgccaaggcc	ttgcagaaca	agccgcgtc	1020
	tccggcttgt	tccaaacctt	caacgtgggt	ggatccggcg	gaggcggcac	ttctgcgccc	1080
	gacttcaatg	caggcgttac	cggtatcgcc	agcaacagca	gagcaacaac	agcggaaatca	1140
	gcagcgttat	cttacggcccg	tatcaagaac	gaaatgttgc	aagacagaaag	cattcttgc	1200
	gccccgtccgg	atgacgttgc	gttacagac	aggatgtcc	aatcaatgc	cccccccccg	1260
	aatctgcata	ccggagactt	tccaaacccaa	aatagcgtat	acaagaattt	gatcaacactc	1320
	aaacctgcaa	ttgaagcagg	ctatacagga	cgccccgtat	aggttaggtat	cgtcgacaca	1380
	ggcgaatccg	tcggcagcat	atccctttcc	gaactgtatg	gcagaaaaaga	acacggctat	1440
	aacggaaatt	acaaaaacta	tacggcgtat	atgcggaaagg	aagcgcctg	agacggaggc	1500
	ggtaaagaca	ttgaagcttc	tttgcacgt	gaggccgtt	tagagactga	agcaaagccg	1560
	acggatatac	gcccacgtaa	agaaatcgg	cacatcgatt	ttgtctccca	tattattggc	1620
	gggcgttccg	tggacggcag	acctgcaggc	gttatttgc	ccgatgcac	gctacacata	1680
	atgaaatacga	atgtatgaaac	caagaacgaa	atgtatgttgc	cagccatccg	caatgcattg	1740
	gtcaagctgg	gcgaaacgtgg	cgtgcgc	gtcaataaca	gttttggaa	aacatcgagg	1800
	gcaggcactg	ccgacccctt	ccaaatagcc	aattcggagg	agcagtaccg	ccaagcgtt	1860
	ctcgactatt	ccggccgttgc	taaaacagac	gagggtatcc	gcctgtatgc	acagagcgt	1920
	tacggcaacc	tgttcttacca	catccgtat	aaaacatgc	ttttcatctt	ttcgacaggg	1980
	aatgcgtcac	aagtcgtacc	caacacat	gcccattatgc	cattttatga	aaaagacgt	2040
	aaaaaggca	ttatcagatgt	cgcggcgta	gaccggcgt	gagaaaatgt	caaaacggaa	2100
	atgtatggag	aacccgggtac	agaaccgtt	gagttatggct	ccaaaccattg	cggaaattact	2160
	gccatgttgt	gcctgtcggc	accctatgaa	gcaagcgtcc	gtttcacccg	tacaaacccg	2220
	attcaatttgc	ccgaaacatc	cttttccgca	cccatcgtaa	ccggcaccggc	ggctctgtct	2280
	ctgcagaaat	accctgttgtat	gagcaacgac	aacctgcgt	ccacgttgc	gacgacggct	2340
	caggacatcg	gtgcgttgc	cgtggac	aatttgcgt	ggggactgt	ggatgcgggt	2400
	aaggccatgt	acggacccgc	gtcccttcc	ttccggcact	ttaccggcga	tacgaaaggt	2460
	acatccgata	ttgccttactc	cttccgtat	gacatttgc	gcacggccgg	cctgtacaaa	2520
	aaaggccgca	gccaactgtc	actgcacggc	aaaaacaccc	atacggccaa	aaccattatc	2580
	gaaggccgtt	cgctgggttt	gtacggcaac	aaaaaatcg	atatgcgt	cgaaacccaa	2640
	ggtgcgtga	tttataacgg	ggcggcatcc	ggccggcagcc	tgaacagcga	cgccattgtc	2700
	tatctggcag	ataccggacca	atccggcgc	aaaaaaaccc	tacacatcaa	aggcagtctg	2760
	cagctggacg	gcaaaaggatc	gctgtacaca	cgtttggca	aactgtctaa	agtggacgg	2820
	acggcgat	tcggcggcaa	gctgtacatg	tcggcacgc	gcaagggggc	aggttatctc	2880
	aacagtaccg	gacgacgtgt	tcccttcc	agtgccgc	aaatcggca	ggattattct	2940
	ttcttcacaa	acatcgaaac	cgacggccgc	ctgtggctt	ccctcgacag	ctgtgaaaaaa	3000
	acagcgggca	gtgaaggcga	cacgctgtcc	tattatgtcc	gtcgcggcaa	tgcggcacgg	3060
	actgtcttcg	cagccggcaca	ttccgcgc	gcccgtctga	aacacggccgt	agaaacagggc	3120
	ggcagcaatc	tggaaaaccc	gatggtgc	ctggatgc	ccgaatcatc	cgcacacc	3180
	gagacgggtt	aaactcgcc	agccgaccgc	acagatatgc	cgggcatcc	ccccacccac	3240
	gcaactttcc	gcccacgtcc	agccgtat	catgcgtat	ccggccgacgg	tgtacgcac	3300
	ttcaacagtc	tcggccgtac	cgtctatgc	gacagtacc	ccgccccatgc	cgatatgcac	3360

	ggacgcccgc	tgaaagccgt	atcggacgg	ttggaccaca	acggcacgg	tctgcgcgtc	3420
	atcgcgaaa	cccaacagga	cggtggaacg	tggaaacagg	gcggtgttga	aggcaaaatg	3480
5	cgcggcagta	cccaaaccgt	cggcattgcc	gcaaaaacccg	gcgaaaatac	gacagcagcc	3540
	gcccacactgg	gcatgggacg	cagcacatgg	agcggaaaaca	tgcaaatgc	aaaaaccgac	3600
	agcatttagtc	tgtttgcagg	catacgcac	gatgcggcg	atatcggtca	tctcaaaggc	3660
	ctgttctcct	acggacgcta	aaaaaacgc	atcagccga	gcaccgggtc	ggacgaacat	3720
	gcggaaaggca	gcgtcaacgg	cacgctgatg	cagctggcg	cactggcg	tgtcaacgtt	3780
	ccgtttgccc	caacgggaga	tttgcgttc	gaaggcggtc	tgcgtacga	cctgctaaa	3840
10	caggatgcat	tcggccaaaa	aggcagtgtc	ttgggctgga	gcggcaacag	cctcaactgaa	3900
	ggcacgcgtg	tgcgactcgc	gggtctgaag	ctgtcgcaac	ccttgagcga	taaagccgtc	3960
	ctgttgcggaa	cggcggggt	ggaacgcgac	ctgaacggac	gcgactacac	gttaacgggc	4020
	ggctttaaccg	gcgcgactgc	agcaaccggc	aagacggggg	cacgcaatat	gccgcacacc	4080
	cgtctgggtt	ccggcctggg	cgggatgtc	gaattcggca	acggctggaa	cggcttggca	4140
	cgttacagct	acggcgggtc	caaacagttac	ggcaaccaca	gcggacgagt	cggcgtaggc	4200
	taccggttct	gactcgag					4218
15	<210>	163					
	<211>	1403					
	<212>	PRT					
	<213>	Artificial Sequence					
20	<220>						
	<223>	961cL-983					
	<400>	163					
	Met Lys His Phe Pro Ser Lys Val Leu Thr Thr Ala Ile Leu Ala Thr						
	1	5	10			15	
25	Phe Cys Ser Gly Ala Leu Ala Ala Thr Asn Asp Asp Asp Val Lys Lys						
	20	25	30				
	Ala Ala Thr Val Ala Ile Ala Ala Ala Tyr Asn Asn Gly Gln Glu Ile						
	35	40	45				
30	Asn Gly Phe Lys Ala Gly Glu Thr Ile Tyr Asp Ile Asp Glu Asp Gly						
	50	55	60				
	Thr Ile Thr Lys Lys Asp Ala Thr Ala Ala Asp Val Glu Ala Asp Asp						
	65	70	75			80	
35	Phe Lys Gly Leu Gly Leu Lys Val Val Thr Asn Leu Thr Lys Thr						
	85	90	95				
	Val Asn Glu Asn Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala Glu						
	100	105	110				
40	Ser Glu Ile Glu Lys Leu Thr Thr Lys Leu Ala Asp Thr Asp Ala Ala						
	115	120	125				
	Leu Ala Asp Thr Asp Ala Ala Leu Asp Ala Thr Thr Asn Ala Leu Asn						
	130	135	140				
45	Lys Leu Gly Glu Asn Ile Thr Thr Phe Ala Glu Glu Thr Lys Thr Asn						
	145	150	155			160	
	Ile Val Lys Ile Asp Glu Lys Leu Glu Ala Val Ala Asp Thr Val Asp						
	165	170	175				
50	Lys His Ala Glu Ala Phe Asn Asp Ile Ala Asp Ser Leu Asp Glu Thr						
	180	185	190				
	Asn Thr Lys Ala Asp Glu Ala Val Lys Thr Ala Asn Glu Ala Lys Gln						
	195	200	205				
	Thr Ala Glu Glu Thr Lys Gln Asn Val Asp Ala Lys Val Lys Ala Ala						
	210	215	220				
55	Glu Thr Ala Ala Gly Lys Ala Glu Ala Ala Ala Gly Thr Ala Asn Thr						
	225	230	235			240	

Ala Ala Asp Lys Ala Glu Ala Val Ala Ala Lys Val Thr Asp Ile Lys  
 245 250 255  
 5 Ala Asp Ile Ala Thr Asn Lys Asp Asn Ile Ala Lys Lys Ala Asn Ser  
 260 265 270  
 Ala Asp Val Tyr Thr Arg Glu Glu Ser Asp Ser Lys Phe Val Arg Ile  
 275 280 285  
 10 Asp Gly Leu Asn Ala Thr Thr Glu Lys Leu Asp Thr Arg Leu Ala Ser  
 290 295 300  
 Ala Glu Lys Ser Ile Ala Asp His Asp Thr Arg Leu Asn Gly Leu Asp  
 305 310 315 320  
 15 Lys Thr Val Ser Asp Leu Arg Lys Glu Thr Arg Gln Gly Leu Ala Glu  
 325 330 335  
 Gln Ala Ala Leu Ser Gly Leu Phe Gln Pro Tyr Asn Val Gly Gly Ser  
 340 345 350  
 20 Gly Gly Gly Thr Ser Ala Pro Asp Phe Asn Ala Gly Gly Thr Gly  
 355 360 365  
 Ile Gly Ser Asn Ser Arg Ala Thr Thr Ala Lys Ser Ala Ala Val Ser  
 370 375 380  
 25 Tyr Ala Gly Ile Lys Asn Glu Met Cys Lys Asp Arg Ser Met Leu Cys  
 385 390 395 400  
 Ala Gly Arg Asp Asp Val Ala Val Thr Asp Arg Asp Ala Lys Ile Asn  
 405 410 415  
 30 Ala Pro Pro Pro Asn Leu His Thr Gly Asp Phe Pro Asn Pro Asn Asp  
 420 425 430  
 Ala Tyr Lys Asn Leu Ile Asn Leu Lys Pro Ala Ile Glu Ala Gly Tyr  
 435 440 445  
 35 Thr Gly Arg Gly Val Glu Val Gly Ile Val Asp Thr Gly Glu Ser Val  
 450 455 460  
 Gly Ser Ile Ser Phe Pro Glu Leu Tyr Gly Arg Lys Glu His Gly Tyr  
 465 470 475 480  
 40 Asn Glu Asn Tyr Lys Asn Tyr Thr Ala Tyr Met Arg Lys Glu Ala Pro  
 485 490 495  
 Glu Asp Gly Gly Lys Asp Ile Glu Ala Ser Phe Asp Asp Glu Ala  
 500 505 510  
 45 Val Ile Glu Thr Glu Ala Lys Pro Thr Asp Ile Arg His Val Lys Glu  
 515 520 525  
 Ile Gly His Ile Asp Leu Val Ser His Ile Ile Gly Gly Arg Ser Val  
 530 535 540  
 Asp Gly Arg Pro Ala Gly Gly Ile Ala Pro Asp Ala Thr Leu His Ile  
 545 550 555 560  
 50 Met Asn Thr Asn Asp Glu Thr Lys Asn Glu Met Met Val Ala Ala Ile  
 565 570 575  
 Arg Asn Ala Trp Val Lys Leu Gly Glu Arg Gly Val Arg Ile Val Asn  
 580 585 590  
 55 Asn Ser Phe Gly Thr Thr Ser Arg Ala Gly Thr Ala Asp Leu Phe Gln  
 595 600 605

## EP 1 790 660 A2

Ile Ala Asn Ser Glu Glu Gln Tyr Arg Gln Ala Leu Leu Asp Tyr Ser  
 610 615 620  
 5 Gly Gly Asp Lys Thr Asp Glu Gly Ile Arg Leu Met Gln Gln Ser Asp  
 625 630 635 640  
 Tyr Gly Asn Leu Ser Tyr His Ile Arg Asn Lys Asn Met Leu Phe Ile  
 645 650 655  
 10 Phe Ser Thr Gly Asn Asp Ala Gln Ala Gln Pro Asn Thr Tyr Ala Leu  
 660 665 670  
 Leu Pro Phe Tyr Glu Lys Asp Ala Gln Lys Gly Ile Ile Thr Val Ala  
 675 680 685  
 15 Gly Val Asp Arg Ser Gly Glu Lys Phe Lys Arg Glu Met Tyr Gly Glu  
 690 695 700  
 Pro Gly Thr Glu Pro Leu Glu Tyr Gly Ser Asn His Cys Gly Ile Thr  
 705 710 715 720  
 20 Ala Met Trp Cys Leu Ser Ala Pro Tyr Glu Ala Ser Val Arg Phe Thr  
 725 730 735  
 Arg Thr Asn Pro Ile Gln Ile Ala Gly Thr Ser Phe Ser Ala Pro Ile  
 740 745 750  
 25 Val Thr Gly Thr Ala Ala Leu Leu Leu Gln Lys Tyr Pro Trp Met Ser  
 755 760 765  
 Asn Asp Asn Leu Arg Thr Thr Leu Leu Thr Thr Ala Gln Asp Ile Gly  
 770 775 780  
 30 Ala Val Gly Val Asp Ser Lys Phe Gly Trp Gly Leu Leu Asp Ala Gly  
 785 790 795 800  
 Lys Ala Met Asn Gly Pro Ala Ser Phe Pro Phe Gly Asp Phe Thr Ala  
 805 810 815  
 35 Asp Thr Lys Gly Thr Ser Asp Ile Ala Tyr Ser Phe Arg Asn Asp Ile  
 820 825 830  
 Ser Gly Thr Gly Leu Ile Lys Lys Gly Ser Gln Leu Gln Leu  
 835 840 845  
 40 His Gly Asn Asn Thr Tyr Thr Gly Lys Thr Ile Ile Glu Gly Ser  
 850 855 860  
 Leu Val Leu Tyr Gly Asn Asn Lys Ser Asp Met Arg Val Glu Thr Lys  
 865 870 875 880  
 45 Gly Ala Leu Ile Tyr Asn Gly Ala Ala Ser Gly Gly Ser Leu Asn Ser  
 885 890 895  
 Asp Gly Ile Val Tyr Leu Ala Asp Thr Asp Gln Ser Gly Ala Asn Glu  
 900 905 910  
 50 Thr Val His Ile Lys Gly Ser Leu Gln Leu Asp Gly Lys Gly Thr Leu  
 915 920 925  
 Tyr Thr Arg Leu Gly Lys Leu Leu Lys Val Asp Gly Thr Ala Ile Ile  
 930 935 940  
 55 Gly Gly Lys Leu Tyr Met Ser Ala Arg Gly Lys Gly Ala Gly Tyr Leu  
 945 950 955 960  
 Asn Ser Thr Gly Arg Arg Val Pro Phe Leu Ser Ala Ala Lys Ile Gly  
 965 970 975

Gln Asp Tyr Ser Phe Phe Thr Asn Ile Glu Thr Asp Gly Gly Leu Leu  
 980 985 990  
 5 Ala Ser Leu Asp Ser Val Glu Lys Thr Ala Gly Ser Glu Gly Asp Thr  
 995 1000 1005  
 Leu Ser Tyr Tyr Val Arg Arg Gly Asn Ala Ala Arg Thr Ala Ser Ala  
 1010 1015 1020  
 10 Ala Ala His Ser Ala Pro Ala Gly Leu Lys His Ala Val Glu Gln Gly  
 1025 1030 1035 1040  
 Gly Ser Asn Leu Glu Asn Leu Met Val Glu Leu Asp Ala Ser Glu Ser  
 1045 1050 1055  
 15 Ser Ala Thr Pro Glu Thr Val Glu Thr Ala Ala Ala Asp Arg Thr Asp  
 1060 1065 1070  
 Met Pro Gly Ile Arg Pro Tyr Gly Ala Thr Phe Arg Ala Ala Ala Ala  
 1075 1080 1085  
 20 Val Gln His Ala Asn Ala Ala Asp Gly Val Arg Ile Phe Asn Ser Leu  
 1090 1095 1100  
 Ala Ala Thr Val Tyr Ala Asp Ser Thr Ala Ala His Ala Asp Met Gln  
 1105 1110 1115 1120  
 25 Gly Arg Arg Leu Lys Ala Val Ser Asp Gly Leu Asp His Asn Gly Thr  
 1125 1130 1135  
 Gly Leu Arg Val Ile Ala Gln Thr Gln Gln Asp Gly Gly Thr Trp Glu  
 1140 1145 1150  
 30 Gln Gly Val Glu Gly Lys Met Arg Gly Ser Thr Gln Thr Val Gly  
 1155 1160 1165  
 Ile Ala Ala Lys Thr Gly Glu Asn Thr Thr Ala Ala Ala Thr Leu Gly  
 1170 1175 1180  
 35 Met Gly Arg Ser Thr Trp Ser Glu Asn Ser Ala Asn Ala Lys Thr Asp  
 1185 1190 1195 1200  
 Ser Ile Ser Leu Phe Ala Gly Ile Arg His Asp Ala Gly Asp Ile Gly  
 1205 1210 1215  
 40 Tyr Leu Lys Gly Leu Phe Ser Tyr Gly Arg Tyr Lys Asn Ser Ile Ser  
 1220 1225 1230  
 Arg Ser Thr Gly Ala Asp Glu His Ala Glu Gly Ser Val Asn Gly Thr  
 1235 1240 1245  
 45 Leu Met Gln Leu Gly Ala Leu Gly Gly Val Asn Val Pro Phe Ala Ala  
 1250 1255 1260  
 Thr Gly Asp Leu Thr Val Glu Gly Gly Leu Arg Tyr Asp Leu Leu Lys  
 1265 1270 1275 1280  
 50 Gln Asp Ala Phe Ala Glu Lys Gly Ser Ala Leu Gly Trp Ser Gly Asn  
 1285 1290 1295  
 Ser Leu Thr Glu Gly Thr Leu Val Gly Leu Ala Gly Leu Lys Leu Ser  
 1300 1305 1310  
 Gln Pro Leu Ser Asp Lys Ala Val Leu Phe Ala Thr Ala Gly Val Glu  
 1315 1320 1325  
 55 Arg Asp Leu Asn Gly Arg Asp Tyr Thr Val Thr Gly Gly Phe Thr Gly  
 1330 1335 1340

	Ala Thr Ala Ala Thr Gly Lys Thr Gly Ala Arg Asn Met Pro His Thr	
	1345 1350 1355 1360	
5	Arg Leu Val Ala Gly Leu Gly Ala Asp Val Glu Phe Gly Asn Gly Trp	
	1365 1370 1375	
	Asn Gly Leu Ala Arg Tyr Ser Tyr Ala Gly Ser Lys Gln Tyr Gly Asn	
	1380 1385 1390	
10	His Ser Gly Arg Val Gly Val Gly Tyr Arg Phe	
	1395 1400	
	<210> 164	
	<211> 33	
	<212> DNA	
15	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 164	
20	cgcggatccg ctagcaaaac aaccgacaaa cgg	33
	<210> 165	
	<211> 27	
	<212> DNA	
	<213> Artificial Sequence	
25	<220>	
	<223> Oligonucleotide	
	<400> 165	
	cccgctcgag ttaccagcgg tagccta	27
30	<210> 166	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
35	<223> Oligonucleotide	
	<400> 166	
	ctagctagcg gacacactta tttcggcatc	30
	<210> 167	
	<211> 32	
40	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
45	<400> 167	
	cccgctcgag ttaccagcgg tagcctaatt tg	32
	<210> 168	
	<211> 10	
	<212> DNA	
50	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 168	
55	cccgctcgag	10
	<210> 169	

	<211>	33	
	<212>	DNA	
	<213>	Artificial Sequence	
5	<220>		
	<223>	Oligonucleotide	
	<400>	169	
		cgcgatccc atatgaaaac cttcttcaaa acc	33
10	<210>	170	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
15	<220>		
	<223>	Oligonucleotide	
	<400>	170	
		cccgctcgag ttatggct gcgccttc	28
20	<210>	171	
	<211>	35	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
25	<400>	171	
		gcggcattaa tatgttgaga aaattgttga aatgg	35
	<210>	172	
	<211>	34	
	<212>	DNA	
30	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	172	
35		gcggcctcga gttatTTTTT caaaatataat ttgc	34
	<210>	173	
	<211>	33	
	<212>	DNA	
	<213>	Artificial Sequence	
40	<220>		
	<223>	Oligonucleotide	
	<400>	173	
		gcggccatat gttacctaac cgTTTcaaaa tgt	33
45	<210>	174	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
50	<223>	Oligonucleotide	
	<400>	174	
		gcggcctcga gttatTTCG aggtttcgg g	31
55	<210>	175	
	<211>	32	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	Oligonucleotide	
5	<400>	175	
	cgccggatccc	atatgacacg cttcaaataat tc	32
	<210>	176	
	<211>	31	
	<212>	DNA	
10	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	176	
	cccgctcgag	ttatTTaaac cgataggtaa a	31
15	<210>	177	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
20	<220>		
	<223>	Oligonucleotide	
	<400>	177	
	cgccggatccc	atatgggcag ggaagaaccg c	31
25	<210>	178	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
30	<220>		
	<223>	Oligonucleotide	
	<400>	178	
	gcccaagctt	atcgatggaa tagccgcg	28
35	<210>	179	
	<211>	32	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
40	<400>	179	
	cgccggatccg	ctagcaacgg tttggatgcc cg	32
	<210>	180	
	<211>	30	
	<212>	DNA	
45	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	180	
50	cccgctcgag	tttgtctaag ttccctgatat	30
	<210>	181	
	<211>	26	
	<212>	DNA	
	<213>	Artificial Sequence	
55	<220>		
	<223>	Oligonucleotide	

	<400> 181	
	<code>cccgctcgag attcccacct gccatc</code>	26
5	<210> 182	
	<211> 37	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> Oligonucleotide	
	<400> 182	
	<code>cgcggatccg ctagcatgaa tttgcctatt caaaaat</code>	37
15	<210> 183	
	<211> 29	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
20	<400> 183	
	<code>cccgctcgag ttaattccca cctgccatc</code>	29
	<210> 184	
	<211> 37	
	<212> DNA	
25	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 184	
30	<code>cgcggatccg ctagcatgaa tttgcctatt caaaaat</code>	37
	<210> 185	
	<211> 27	
	<212> DNA	
	<213> Artificial Sequence	
35	<220>	
	<223> Oligonucleotide	
	<400> 185	
	<code>cccgctcgag ttggacgtt cccgcga</code>	27
40	<210> 186	
	<211> 37	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
45	<223> Oligonucleotide	
	<400> 186	
	<code>cgcggatccg ctagcatgaa tttgcctatt caaaaat</code>	37
50	<210> 187	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
55	<400> 187	
	<code>cccgctcgag ttattggacg atgcccgc</code>	28

	<210> 188	
5	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
10	<400> 188	
	cgcgatccc atatgtatcg caaactgatt gc	32
	<210> 189	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
15	<220>	
	<223> Oligonucleotide	
	<400> 189	
	cccgctcgag ctaatcgatg gaatagcc	28
20	<210> 190	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
25	<223> Oligonucleotide	
	<400> 190	
	cgcgatccc atatgaaaca gacagtcaaa tg	32
30	<210> 191	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
35	<400> 191	
	cccgctcgag tcaataaccc gccttcag	28
	<210> 192	
	<211> 45	
	<212> DNA	
40	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 192	
45	cgcgatccc atatgttacg tttgactgct ttagccgtat gcacc	45
	<210> 193	
	<211> 40	
	<212> DNA	
	<213> Artificial Sequence	
50	<220>	
	<223> Oligonucleotide	
	<400> 193	
	cccgctcgag ttatTTTgCC gcgttaaaag cgTCGGcaac	40
55	<210> 194	
	<211> 32	

<212>	DNA		
<213>	Artificial Sequence		
5			
<220>			
<223>	Oligonucleotide		
<400>	194		
	cgcgatccc atatgaacaa aataccgc at		32
10			
<210>	195		
<211>	28		
<212>	DNA		
<213>	Artificial Sequence		
<220>			
<223>	Oligonucleotide		
15			
<400>	195		
	cccgctcgag ttaccactga taaccgac		28
20			
<210>	196		
<211>	34		
<212>	DNA		
<213>	Artificial Sequence		
<220>			
<223>	Oligonucleotide		
25			
<400>	196		
	cgcgatccc atatgaccga tgacgacgat ttat		34
30			
<210>	197		
<211>	28		
<212>	DNA		
<213>	Artificial Sequence		
<220>			
<223>	Oligonucleotide		
35			
<400>	197		
	ccccaaagtt ccactgataa ccgacaga		28
40			
<210>	198		
<211>	32		
<212>	DNA		
<213>	Artificial Sequence		
<220>			
<223>	Oligonucleotide		
45			
<400>	198		
	cgcgatccc atatgaacaa aataccgc at		32
50			
<210>	199		
<211>	28		
<212>	DNA		
<213>	Artificial Sequence		
<220>			
<223>	Oligonucleotide		
55			
<400>	199		
	ccccaaagtt ttaccactga taaccgac		28
<210>	200		
<211>	34		
<212>	DNA		
<213>	Artificial Sequence		

<220>		
<223>	Oligonucleotide	
5	<400> 200	
	gggaattcca tatgggcatt tcccgcaaaa tatac	34
	<210> 201	
	<211> 40	
	<212> DNA	
10	<213> Artificial Sequence	
	<220>	
	<223>	Oligonucleotide
	<400> 201	
	cccgctcgag ttatttactc ctataacgag gtctcttaac	40
15	<210> 202	
	<211> 36	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223>	Oligonucleotide
	<400> 202	
	gggaattcca tatgtcagat ttggcaaacg attctt	36
25	<210> 203	
	<211> 40	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223>	Oligonucleotide
	<400> 203	
	cccgctcgag ttatttactc ctataacgag gtctcttaac	40
35	<210> 204	
	<211> 34	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223>	Oligonucleotide
40	<400> 204	
	gggaattcca tatgggcatt tcccgcaaaa tatac	34
	<210> 205	
	<211> 32	
	<212> DNA	
45	<213> Artificial Sequence	
	<220>	
	<223>	Oligonucleotide
	<400> 205	
50	cccgctcgag ttacgtatca tatttcacgt gc	32
	<210> 206	
	<211> 35	
	<212> DNA	
	<213> Artificial Sequence	
55	<220>	
	<223>	Oligonucleotide

	<400> 206		
	gggaattcca tatgcacgtg aaatatgata cgaag		35
5	<210> 207		
	<211> 37		
	<212> DNA		
	<213> Artificial Sequence		
10	<220>		
	<223> Oligonucleotide		
	<400> 207		
	cccgctcgag tttactccta taacgaggtc tcttaac		37
15	<210> 208		
	<211> 36		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
	<223> Oligonucleotide		
20	<400> 208		
	gggaattcca tatgtcagat ttggcaaacg attctt		36
	<210> 209		
	<211> 29		
	<212> DNA		
25	<213> Artificial Sequence		
	<220>		
	<223> Oligonucleotide		
	<400> 209		
30	cccgctcgag cgtatcatat ttcacgtgc		29
	<210> 210		
	<211> 36		
	<212> DNA		
	<213> Artificial Sequence		
35	<220>		
	<223> Oligonucleotide		
	<400> 210		
	gggaattcca tatgtcagat ttggcaaacg attctt		36
40	<210> 211		
	<211> 37		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
45	<223> Oligonucleotide		
	<400> 211		
	cccgctcgag tttactccta taacgaggtc tcttaac		37
50	<210> 212		
	<211> 35		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
	<223> Oligonucleotide		
55	<400> 212		
	cgcggatccc atatgcaaaa tgcgttcaaa atccc		35

	<210>	213																																																																																																																																																																																																					
	<211>	32																																																																																																																																																																																																					
	<212>	DNA																																																																																																																																																																																																					
5	<213>	Artificial Sequence																																																																																																																																																																																																					
	<220>																																																																																																																																																																																																						
	<223>	Oligonucleotide																																																																																																																																																																																																					
	<400>	213																																																																																																																																																																																																					
	cgcgatccc	atatgaacaa	aatataccgc	10			32		<210>	214			<211>	29			<212>	DNA			<213>	Artificial Sequence		15	<220>				<223>	Oligonucleotide			<400>	214			cccgctcgag	tttgcttcg	atagaacgg	20			29		<210>	215			<211>	34			<212>	DNA			<213>	Artificial Sequence			<220>			25	<223>	Oligonucleotide			<400>	215			gcggccatat	ggtcataaaa	tatacaaatt	30			34		<210>	216			<211>	34			<212>	DNA			<213>	Artificial Sequence			<220>				<223>	Oligonucleotide		35	<400>	216			gcggcctcga	gttagcctga	gaccttgca				34		<210>	217			<211>	30			<212>	DNA		40	<213>	Artificial Sequence			<220>				<223>	Oligonucleotide			<400>	217			gcggccatat	gaaacagaaa	aaaaccgctg	45			30		<210>	218			<211>	32			<212>	DNA			<213>	Artificial Sequence		50	<220>				<223>	Oligonucleotide			<400>	218			gcggcctcga	gttacggttt	gacaccgttt	55			32		<210>	219			<211>	30			<212>	DNA	
10			32																																																																																																																																																																																																				
	<210>	214																																																																																																																																																																																																					
	<211>	29																																																																																																																																																																																																					
	<212>	DNA																																																																																																																																																																																																					
	<213>	Artificial Sequence																																																																																																																																																																																																					
15	<220>																																																																																																																																																																																																						
	<223>	Oligonucleotide																																																																																																																																																																																																					
	<400>	214																																																																																																																																																																																																					
	cccgctcgag	tttgcttcg	atagaacgg	20			29		<210>	215			<211>	34			<212>	DNA			<213>	Artificial Sequence			<220>			25	<223>	Oligonucleotide			<400>	215			gcggccatat	ggtcataaaa	tatacaaatt	30			34		<210>	216			<211>	34			<212>	DNA			<213>	Artificial Sequence			<220>				<223>	Oligonucleotide		35	<400>	216			gcggcctcga	gttagcctga	gaccttgca				34		<210>	217			<211>	30			<212>	DNA		40	<213>	Artificial Sequence			<220>				<223>	Oligonucleotide			<400>	217			gcggccatat	gaaacagaaa	aaaaccgctg	45			30		<210>	218			<211>	32			<212>	DNA			<213>	Artificial Sequence		50	<220>				<223>	Oligonucleotide			<400>	218			gcggcctcga	gttacggttt	gacaccgttt	55			32		<210>	219			<211>	30			<212>	DNA																																					
20			29																																																																																																																																																																																																				
	<210>	215																																																																																																																																																																																																					
	<211>	34																																																																																																																																																																																																					
	<212>	DNA																																																																																																																																																																																																					
	<213>	Artificial Sequence																																																																																																																																																																																																					
	<220>																																																																																																																																																																																																						
25	<223>	Oligonucleotide																																																																																																																																																																																																					
	<400>	215																																																																																																																																																																																																					
	gcggccatat	ggtcataaaa	tatacaaatt	30			34		<210>	216			<211>	34			<212>	DNA			<213>	Artificial Sequence			<220>				<223>	Oligonucleotide		35	<400>	216			gcggcctcga	gttagcctga	gaccttgca				34		<210>	217			<211>	30			<212>	DNA		40	<213>	Artificial Sequence			<220>				<223>	Oligonucleotide			<400>	217			gcggccatat	gaaacagaaa	aaaaccgctg	45			30		<210>	218			<211>	32			<212>	DNA			<213>	Artificial Sequence		50	<220>				<223>	Oligonucleotide			<400>	218			gcggcctcga	gttacggttt	gacaccgttt	55			32		<210>	219			<211>	30			<212>	DNA																																																																									
30			34																																																																																																																																																																																																				
	<210>	216																																																																																																																																																																																																					
	<211>	34																																																																																																																																																																																																					
	<212>	DNA																																																																																																																																																																																																					
	<213>	Artificial Sequence																																																																																																																																																																																																					
	<220>																																																																																																																																																																																																						
	<223>	Oligonucleotide																																																																																																																																																																																																					
35	<400>	216																																																																																																																																																																																																					
	gcggcctcga	gttagcctga	gaccttgca				34		<210>	217			<211>	30			<212>	DNA		40	<213>	Artificial Sequence			<220>				<223>	Oligonucleotide			<400>	217			gcggccatat	gaaacagaaa	aaaaccgctg	45			30		<210>	218			<211>	32			<212>	DNA			<213>	Artificial Sequence		50	<220>				<223>	Oligonucleotide			<400>	218			gcggcctcga	gttacggttt	gacaccgttt	55			32		<210>	219			<211>	30			<212>	DNA																																																																																																													
			34																																																																																																																																																																																																				
	<210>	217																																																																																																																																																																																																					
	<211>	30																																																																																																																																																																																																					
	<212>	DNA																																																																																																																																																																																																					
40	<213>	Artificial Sequence																																																																																																																																																																																																					
	<220>																																																																																																																																																																																																						
	<223>	Oligonucleotide																																																																																																																																																																																																					
	<400>	217																																																																																																																																																																																																					
	gcggccatat	gaaacagaaa	aaaaccgctg	45			30		<210>	218			<211>	32			<212>	DNA			<213>	Artificial Sequence		50	<220>				<223>	Oligonucleotide			<400>	218			gcggcctcga	gttacggttt	gacaccgttt	55			32		<210>	219			<211>	30			<212>	DNA																																																																																																																																																	
45			30																																																																																																																																																																																																				
	<210>	218																																																																																																																																																																																																					
	<211>	32																																																																																																																																																																																																					
	<212>	DNA																																																																																																																																																																																																					
	<213>	Artificial Sequence																																																																																																																																																																																																					
50	<220>																																																																																																																																																																																																						
	<223>	Oligonucleotide																																																																																																																																																																																																					
	<400>	218																																																																																																																																																																																																					
	gcggcctcga	gttacggttt	gacaccgttt	55			32		<210>	219			<211>	30			<212>	DNA																																																																																																																																																																																					
55			32																																																																																																																																																																																																				
	<210>	219																																																																																																																																																																																																					
	<211>	30																																																																																																																																																																																																					
	<212>	DNA																																																																																																																																																																																																					

	<213> Artificial sequence	
5	<220>	
	<223> oligonucleotide	
	<400> 219	
	cgcgatccc atatgaaaac cctgctcctc	30
10	<210> 220	
	<211> 27	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
15	<400> 220	
	cccgctcgag ttatcctcct ttgcggc	27
	<210> 221	
	<211> 30	
	<212> DNA	
20	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
	<400> 221	
25	gcggccatat ggcaaaatg atgaaatggg	30
	<210> 222	
	<211> 29	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> oligonucleotide	
	<400> 222	
	gcggccctcga gttatcgcg cgccgggcc	29
35	<210> 223	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
40	<223> oligonucleotide	
	<400> 223	
	gcggccatat gaaaaaatcc tccctcatca	30
	<210> 224	
	<211> 32	
	<212> DNA	
45	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
50	<400> 224	
	gcggccctcga gttatttgcc gccgttttg gc	32
	<210> 225	
	<211> 31	
	<212> DNA	
55	<213> Artificial Sequence	
	<220>	

	<223> Oligonucleotide	
5	<400> 225 gcggccatat ggccccctgcc gacgcggtaa g	31
	<210> 226	
	<211> 33	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> Oligonucleotide	
	<400> 226 gcggcctcga gtttgcgc gtttttggct ttc	33
15	<210> 227	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223> Oligonucleotide	
	<400> 227 gcggccatat gaaacacata ctccccctga	30
25	<210> 228	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
30	<400> 228 gcggcctcga gttattcgcc tacggtttt tg	32
	<210> 229	
	<211> 30	
	<212> DNA	
35	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 229 gcggccatat gatttacatc gtactgttcc	30
40	<210> 230	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
45	<220>	
	<223> Oligonucleotide	
	<400> 230 gcggcctcga gttaggagaa caggcgcaat gc	32
50	<210> 231	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
55	<220>	
	<223> Oligonucleotide	
	<400> 231	

	gcggccatat gtacaacatg tatcaggaaa ac	32
5	<210> 232	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
10	<223> Oligonucleotide	
	<400> 232	
	gcggccctcga gggagaacag gcgcaatgcg g	31
	<210> 233	
	<211> 29	
	<212> DNA	
15	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 233	
20	cgcggatccg ctagctgcgg cacggcggg	29
	<210> 234	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
25	<220>	
	<223> Oligonucleotide	
	<400> 234	
	cccgctcgag ataacggtat gccgccag	28
30	<210> 235	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
35	<223> Oligonucleotide	
	<400> 235	
	cgcggatccc atatgaaatc aacactttca c	31
	<210> 236	
40	<211> 27	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
45	<400> 236	
	cccgctcgag ttacacgcgg ttgctgt	27
	<210> 237	
	<211> 31	
	<212> DNA	
50	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 237	
55	cgcggatccc atatgaacaa cagacattt g	31
	<210> 238	

	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
5	<220>		
	<223>	Oligonucleotide	
	<400>	238	
		cccgctcgag ttacctgtcc ggtaaaag	28
10	<210>	239	
	<211>	33	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
15	<223>	Oligonucleotide	
	<400>	239	
		cgcggatccg ctagcaccgt catcaaacag gaa	33
20	<210>	240	
	<211>	27	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
25	<223>	Oligonucleotide	
	<400>	240	
		cccgctcgag tcaagattcg acggggga	27
	<210>	241	
	<211>	31	
30	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	241	
35		cgcggatccc atatgtccgc aaacgaatac g	31
	<210>	242	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
40	<220>		
	<223>	Oligonucleotide	
	<400>	242	
		cccgctcgag tcagtgttct gccagttt	28
45	<210>	243	
	<211>	29	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
50	<223>	Oligonucleotide	
	<400>	243	
		cgcggatccc atatgccgtc tgaaacacg	29
55	<210>	244	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	Oligonucleotide	
5	<400>	244	
	cccgctcgag	tttagcggagc	agtttttc
	<210>	245	
	<211>	28	
	<212>	DNA	
10	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
15	<400>	245	
	cgcggatccc	atatgaccgc	catcagcc
	<210>	246	
	<211>	27	
	<212>	DNA	
20	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
25	<400>	246	
	cccgctcgag	ttaaagccgg	gtaacgc
	<210>	247	
	<211>	31	
	<212>	DNA	
30	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
35	<400>	247	
	gcggccatat	ggaaacacag	ctttacatcg
	<210>	248	
	<211>	30	
	<212>	DNA	
40	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
45	<400>	248	
	gcggcctcga	gtcaataata	atatcccgcg
	<210>	249	
	<211>	30	
	<212>	DNA	
50	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
55	<400>	249	
	gcggccatat	gattaaaatc	cgcaatatcc
	<210>	250	
	<211>	36	
	<212>	DNA	
55	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	

	<400> 250 gcggcctcga gttaaatctt ggttagattgg atttgg	36
5	<210> 251 <211> 30 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide	
15	<400> 251 gcggccatat gactgacaac gcactgctcc	30
20	<210> 252 <211> 31 <212> DNA <213> Artificial Sequence	
25	<220> <223> Oligonucleotide	
30	<400> 252 gcggcctcga gtcagaccgc gttgtcgaaa c	31
35	<210> 253 <211> 32 <212> DNA <213> Artificial Sequence	
40	<220> <223> Oligonucleotide	
45	<400> 253 cgcggatccc atatggcggtt aaaaacatca aa	32
50	<210> 254 <211> 27 <212> DNA <213> Artificial Sequence	
55	<220> <223> Oligonucleotide	
60	<400> 254 cccgctcgag tcagcccttc atacagc	27
65	<210> 255 <211> 32 <212> DNA <213> Artificial Sequence	
70	<220> <223> Oligonucleotide	
75	<400> 255 gcggcattaa tggcacaaac tacactcaa cc	32
80	<210> 256 <211> 33 <212> DNA <213> Artificial Sequence	
85	<220> <223> Oligonucleotide	
90	<400> 256 gcggcctcga gttaaaaactt cacgttcacg ccg	33

	<210>	257			
5	<211>	34			
	<212>	DNA			
	<213>	Artificial Sequence			
	<220>				
	<223>	Oligonucleotide			
	<400>	257			
10	gcggcattaa	tgcataaacc	tgagcaatcg	gtgg	34
	<210>	258			
	<211>	38			
	<212>	DNA			
	<213>	Artificial Sequence			
15	<220>				
	<223>	Oligonucleotide			
	<400>	258			
	gcggcctcga	gaaacttcac	gttcacgccc	ccggtaaa	38
20	<210>	259			
	<211>	33			
	<212>	DNA			
	<213>	Artificial Sequence			
	<220>				
25	<223>	Oligonucleotide			
	<400>	259			
	cgcggatccc	atatggcaa	atccgaaaat	acg	33
30	<210>	260			
	<211>	26			
	<212>	DNA			
	<213>	Artificial Sequence			
	<220>				
	<223>	Oligonucleotide			
35	<400>	260			
	cccgctcgag	ataatggcgg	cggcgg	26	
	<210>	261			
	<211>	29			
	<212>	DNA			
40	<213>	Artificial Sequence			
	<220>				
	<223>	Oligonucleotide			
	<400>	261			
45	cgcggatccc	atatgtttcc	ccccgacaa	29	
	<210>	262			
	<211>	31			
	<212>	DNA			
	<213>	Artificial Sequence			
50	<220>				
	<223>	Oligonucleotide			
	<400>	262			
	cccgctcgag	tcattctgta	aaaaaaagtat	g	31
55	<210>	263			
	<211>	32			

	<212> DNA	
	<213> Artificial Sequence	
5	<220>	
	<223> Oligonucleotide	
	<400> 263	
	cgcggatccc atatgcttca aagcgacagc ag	32
10	<210> 264	
	<211> 29	
	<212> DNA	
	<213> Artificial Sequence	
15	<220>	
	<223> Oligonucleotide	
	<400> 264	
	cccgctcgag ttcggatttt tgctgtactc	29
20	<210> 265	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
25	<400> 265	
	cgcggatccc atatggcaat ggcagaaaaac g	31
	<210> 266	
	<211> 27	
	<212> DNA	
30	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 266	
35	cccgctcgag ctataacaatc cgtgccg	27
	<210> 267	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
40	<220>	
	<223> Oligonucleotide	
	<400> 267	
	cgcggatccc atatggattc ttttttcaaa cc	32
45	<210> 268	
	<211> 27	
	<212> DNA	
	<213> Artificial Sequence	
50	<220>	
	<223> Oligonucleotide	
	<400> 268	
	cccgctcgag tcagttcaga aagcggg	27
55	<210> 269	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	

	<220>		
	<223>	Oligonucleotide	
5	<400>	269	
	cgcgatccc atatgaaacc tttgattta gg		32
	<210>	270	
	<211>	28	
10	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
15	<400>	270	
	cccgctcgag ttatgggc tgctcttc		28
	<210>	271	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
20	<220>		
	<223>	Oligonucleotide	
	<400>	271	
	cgcggatccc atatggtaat cgtctggttt		30
25	<210>	272	
	<211>	27	
	<212>	DNA	
	<213>	Artificial Sequence	
30	<220>		
	<223>	Oligonucleotide	
	<400>	272	
	cccgctcgag ctacgacttg gttaccg		27
35	<210>	273	
	<211>	33	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
40	<400>	273	
	gcggccatat gagacgtaaa atgctaaagc tac		33
	<210>	274	
	<211>	31	
45	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
50	<400>	274	
	gcggccatcgatgtcgtttgcgc c		31
	<210>	275	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
55	<220>		
	<223>	Oligonucleotide	

	<400> 275 gcccgcataat gttgacttta acccgaaaaaa	30
5	<210> 276 <211> 34 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide	
	<400> 276 gccccctcga gcgcggcggt caataccgcc cgaa	34
15	<210> 277 <211> 32 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
20	<400> 277 cgcggatccc atatggcgca atgcgatttg ac	32
25	<210> 278 <211> 27 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
30	<400> 278 cccgctcgag.ttcggcggt aatgccg	27
	<210> 279 <211> 28 <212> DNA <213> Artificial Sequence	
35	<220> <223> Oligonucleotide	
	<400> 279 gcggccataat ggcccccccg atttttgt	28
40	<210> 280 <211> 33 <212> DNA <213> Artificial Sequence	
45	<220> <223> Oligonucleotide	
	<400> 280 gcggccataat gttatggct ttcagtatta ttg	33
50	<210> 281 <211> 30 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
55	<400> 281 gcggccataat gaactttgct ttatccgtca	30

5	<210> 282 <211> 32 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide  <400> 282 gcggcctcga gttaacggca gtatgttt ac	32
15	<210> 283 <211> 31 <212> DNA <213> Artificial Sequence	
20	<220> <223> Oligonucleotide  <400> 283 cgcgatccc atatgggtt ggcgttcggg c	31
25	<210> 284 <211> 29 <212> DNA <213> Artificial Sequence	
30	<220> <223> Oligonucleotide  <400> 284 gccaagtt tttcccttg ccgttccg	29
35	<210> 285 <211> 32 <212> DNA <213> Artificial Sequence	
40	<220> <223> Oligonucleotide  <400> 285 cgcgatccc atatggccga ccttccgaa aa	32
45	<210> 286 <211> 27 <212> DNA <213> Artificial Sequence	
50	<220> <223> Oligonucleotide  <400> 286 cccgctcgag gaagcgcggtt cccaaagc	27
55	<210> 287 <211> 29 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide  <400> 287 cgcgatccc atatgcacga cacccgtac	29
	<210> 288 <211> 28 <212> DNA	

	<213> Artificial Sequence	
5	<220>	
	<223> Oligonucleotide	
	<400> 288	
	cccgctcgag tttagaagcgc gttcccaa	28
10	<210> 289	
	<211> 34	
	<212> DNA	
	<213> Artificial Sequence	.
	<220>	
	<223> Oligonucleotide	
15	<400> 289	
	ctagcttagct ttaaacgcag cgtaatcgca atgg	34
	<210> 290	
	<211> 31	
	<212> DNA	
20	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotideucleotide	
25	<400> 290	
	cccgctcgag tcaatcctgc tctttttgc c	31
	<210> 291	
	<211> 25	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> Oligonucleotide	
	<400> 291	
	ctagcttagcg ggggcggcgg tggcg	25
35	<210> 292	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
40	<220>	
	<223> oligonucleotide	
	<400> 292	
	cccgctcgag tcaatcctgc tctttttgc c	31
	<210> 293	
	<211> 40	
	<212> DNA	
45	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
50	<400> 293	
	ctagcttagcg ctcatcctcg ccgcctgcgg gggcggcgg	40
	<210> 294	
	<211> 31	
	<212> DNA	
55	<213> Artificial Sequence	
	<220>	

	<223> oligonucleotide	
5	<400> 294 cccgctcgag tcaatcctgc tctttttgc c	31
	<210> 295	
	<211> 25	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> Oligonucleotide	
	<400> 295 cggggatccg ggggcggcgg tggcg	25
15	<210> 296	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223> Oligonucleotide	
	<400> 296 cccgctcgag tcaatcctgc tctttttgc c	31
25	<210> 297	
	<211> 25	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
30	<400> 297 ctagctagcg ggggcggcgg tggcg	25
	<210> 298	
	<211> 28	
	<212> DNA	
35	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 298 cccgctcgag atcctgctct ttttgcc	28
40	<210> 299	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
45	<220>	
	<223> Oligonucleotide	
	<400> 299 ctagctagct gcgggggcgg cggtggcg	28
50	<210> 300	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
55	<220>	
	<223> Oligonucleotide	
	<400> 300	

	<b>cccgctcgag atcctgctct ttttgcc</b>	<b>28</b>
5	<210> 301	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
10	<400> 301	
	<b>cgcggatccg ctagccccga tggtaaatcg gc</b>	<b>32</b>
	<210> 302	
	<211> 32	
	<212> DNA	
15	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 302	
20	<b>cgcggatccg ctagccaaga tatggcggca gt</b>	<b>32</b>
	<210> 303	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
25	<220>	
	<223> Oligonucleotide	
	<400> 303	
	<b>cgcggatccg ctagcgccga atccgcaa at ca</b>	<b>32</b>
30	<210> 304	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
35	<223> oligonucleotide	
	<400> 304	
	<b>cgcgctagcg gaagggttga tttggcta at gg</b>	<b>32</b>
	<210> 305	
	<211> 32	
	<212> DNA	
40	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
	<400> 305	
45	<b>cgcgctagcg gaagggttga tttggcta at gg</b>	<b>32</b>
	<210> 306	
	<211> 29	
	<212> DNA	
50	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
	<400> 306	
55	<b>cgcctatgt taaaacgcag cgtaatcgc</b>	<b>29</b>
	<210> 307	

	<211>	34	
	<212>	DNA	
	<213>	Artificial Sequence	
5	<220>		
	<223>	oligonucleotide	
	<400>	307	
		cccgctcgag aaaattgcta ccgccattcg cagg	34
10	<210>	308	
	<211>	32	
	<212>	DNA	
	<213>	Artificial Sequence	
15	<220>		
	<223>	Oligonucleotide	
	<400>	308	
		cggccatatgg gaagggttga tttggctaat gg	32
20	<210>	309	
	<211>	38	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
25	<400>	309	
		cccgctcgag cttgtcttta taaatgtatga catatttg	38
	<210>	310	
	<211>	40	
	<212>	DNA	
30	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	310	
35		cccgctcgag ttataaaaag ataatatattt gattgattcc	40
	<210>	311	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
40	<220>		
	<223>	Oligonucleotide	
	<400>	311	
		cgcgctagca tgccgctgat tccccgtcaat c	31
45	<210>	312	
	<211>	25	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
50	<223>	Oligonucleotide	
	<400>	312	
		ctagcttagcg ggggcggcgg tggcg	25
55	<210>	313	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	oligonucleotide	
5	<400>	313	
	cccgctcgag tcaatcctgc tctttttgc c		31
	<210>	314	
	<211>	32	
	<212>	DNA	
10	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	314	
15	cgcgatccg cttagccccga tgttaaatcg gc		32
	<210>	315	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
20	<220>		
	<223>	oligonucleotide	
	<400>	315	
	cccgctcgag atccctgctct tttttgcc		28
25	<210>	316	
	<211>	32	
	<212>	DNA	
	<213>	Artificial Sequence	
30	<220>		
	<223>	oligonucleotide	
	<400>	316	
	cgcgatccg cttagccccga tgttaaatcg gc		32
35	<210>	317	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
40	<400>	317	
	cccgctcgag tcaatcctgc tctttttgc c		31
	<210>	318	
	<211>	87	
	<212>	DNA	
45	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	318	
50	cgcgatccg cttagtttga acgcagtgtg attgcaatgg cttgtatccc tgccctttca		60
	gcctgtcgc ccgatgttaa atcggcgc		87
	<210>	319	
	<211>	31	
	<212>	DNA	
55	<213>	Artificial Sequence	
	<220>		

	<223> oligonucleotide	
5	<400> 319 cccgctcgag tcaatcctgc tctttttgc c	31
	<210> 320	
	<211> 90	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> oligonucleotide	
	<400> 320 cgcggatccg ctagaaaaac cttcttcaa acccttccg ccggccact cgcgctcatc ctcgccgcct gtcgcccga tgttaaatcg	60 90
15	<210> 321	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223> oligonucleotide	
	<400> 321 cccgctcgag tcaatcctgc tctttttgc c	31
25	<210> 322	
	<211> 33	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> oligonucleotide	
	<400> 322 cgcggatccc atatgaaaac caagttaatc aaa	33
35	<210> 323	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
40	<400> 323 cccgctcgag ttattgattt ttgcggatga	30
	<210> 324	
	<211> 34	
	<212> DNA	
	<213> Artificial Sequence	
45	<220>	
	<223> oligonucleotide	
	<400> 324 cgcggatccc atatgtaaa tcgggtatcc tatac	34
50	<210> 325	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
55	<220>	
	<223> oligonucleotide	

	<400> 325 cccgctcgag ttaatccgcc attccctg	28
5	<210> 326 <211> 30 <212> DNA <213> Artificial Sequence	
10	<220> <223> oligonucleotide	
15	<400> 326 gcggccatat gaaattacaa caattggctg	30
	<210> 327 <211> 31 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
20	<400> 327 gcggcctcga gttaccttac gtttttcaaa g	31
	<210> 328 <211> 29 <212> DNA <213> Artificial Sequence	
25	<220> <223> oligonucleotide	
30	<400> 328 cgcggatccc atatgcaagc acggctgtct	29
	<210> 329 <211> 29 <212> DNA <213> Artificial Sequence	
35	<220> <223> oligonucleotide	
	<400> 329 cccgctcgag tcaagggttgt ccttgctca	29
40	<210> 330 <211> 30 <212> DNA <213> Artificial Sequence	
45	<220> <223> Oligonucleotide	
50	<400> 330 cgcggatccc atatgatgaa accgcacaac	30
	<210> 331 <211> 28 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
55	<400> 331 cccgctcgag tcagttgctc aacacgtc	28

5	<210> 332 <211> 32 <212> DNA <213> Artificial Sequence .	
10	<220> <223> Oligonucleotide  <400> 332 cgcggatccc atatggtaga cgcgcttaag ca	32
15	<210> 333 <211> 25 <212> DNA <213> Artificial sequence	
20	<220> <223> Oligonucleotide  <400> 333 cccgctcgag agctgcatgg cggcg	25
25	<210> 334 <211> 30 <212> DNA <213> Artificial Sequence	
30	<220> <223> Oligonucleotide  <400> 334 cgcggatccc atatggcacg gtcgttatac	30
35	<210> 335 <211> 26 <212> DNA <213> Artificial Sequence	
40	<220> <223> Oligonucleotide  <400> 335 cccgctcgag ctaccgcgca ttcctg	26
45	<210> 336 <211> 31 <212> DNA <213> Artificial Sequence	
50	<220> <223> Oligonucleotide  <400> 336 gcggccatat ggaatttttc attatcttgt t	31
55	<210> 337 <211> 31 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide  <400> 337 gcggccatcgaa gttatggc ggtttgctg c	31
	<210> 338 <211> 32 <212> DNA	

	<213> Artificial Sequence	
5	<220>	
	<223> Oligonucleotide	
	<400> 338	
	gcggccatat gaagtatgtc cggttatttt tc	32
10	<210> 339	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
15	<220>	
	<223> Oligonucleotide	
	<400> 339	
	gcggcctcga gttatcggtc tgtgcaacgg	30
20	<210> 340	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
25	<220>	
	<223> Oligonucleotide	
	<400> 340	
	cgcggatccg ctagctccgg cagcaaaacc ga	32
	<210> 341	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> Oligonucleotide	
	<400> 341	
	gcccaagctt acgcagttcg gaatggag	28
35	<210> 342	
	<211> 35	
	<212> DNA	
	<213> Artificial Sequence	
40	<220>	
	<223> Oligonucleotide	
	<400> 342	
	gccggccatat gttaatattt aaactgaaaa ctttg	35
45	<210> 343	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
50	<220>	
	<223> Oligonucleotide	
	<400> 343	
	gccggcctcga gttatattctg atgccttttc cc	32
55	<210> 344	
	<211> 29	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	

	<223> Oligonucleotide	
5	<400> 344 gccgcctat ggacaataag accaaactg	29
	<210> 345	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> Oligonucleotide	
	<400> 345 gcccgcgtcgaa gttaacggtg cggacgttcc	30
15	<210> 346	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223> Oligonucleotide	
	<400> 346 cgcggatccc atatgaacaa actgtttctt ac	32
25	<210> 347	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
30	<400> 347 cccgctcgag tcattccgccc ttcagaaaa	28
	<210> 348	
	<211> 45	
	<212> DNA	
35	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 348 cgcggatccc atatgcaagg tatcggtgcc gacaaatccg cacct	45
40	<210> 349	
	<211> 42	
	<212> DNA	
	<213> Artificial Sequence	
45	<220>	
	<223> Oligonucleotide	
	<400> 349 cccgctcgag agctaattgt gcttggtttg cagataggag tt	42
50	<210> 350	
	<211> 52	
	<212> DNA	
	<213> Artificial Sequence	
55	<220>	
	<223> Oligonucleotide	
	<400> 350	

	cgcgatccc atatgaaccg caccctgtac aaagttgtat ttaacaaaca tc	52
5	<210> 351	
	<211> 45	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
10	<223> oligonucleotide	
	<400> 351	
	cccgctcgag ttaagctaat tgtgcttggt ttgcagatag gagtt	45
	<210> 352	
	<211> 46	
	<212> DNA	
15	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 352	
20	cgcgatccc atatgacggg agaaaatcat gcggtttcac ttcatg	46
	<210> 353	
	<211> 42	
	<212> DNA	
	<213> Artificial Sequence	
25	<220>	
	<223> Oligonucleotide	
	<400> 353	
	cccgctcgag agctaattgt gcttggtttg cagataggag tt	42
30	<210> 354	
	<211> 51	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
35	<223> Oligonucleotide	
	<400> 354	
	cgcggatccc atatggtttc agacggccta tacaaccaac atggtaaat t	51
	<210> 355	
40	<211> 41	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
45	<400> 355	
	cccgctcgag gcggtaactg ccgcttgcac tgaatccgta a	41
	<210> 356	
	<211> 46	
	<212> DNA	
50	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 356	
55	cgcggatccc atatgacggg agaaaatcat gcggtttcac ttcatg	46
	<210> 357	

	<211>	41	
	<212>	DNA	
	<213>	Artificial sequence	
5	<220>		
	<223>	Oligonucleotide	
	<400>	357	
		cccgctcgag gcggtaactg ccgttgcac tgaatccgta a	41
10	<210>	358	
	<211>	49	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
15	<223>	Oligonucleotide	
	<400>	358	
		cgcggatccc atatgcaaag caaagtcaaa gcagaccatg cctccgtaa	49
20	<210>	359	
	<211>	56	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
25	<223>	Oligonucleotide	
	<400>	359	
		cccgctcgag tctttcctt tcaattataa ctttagtagg ttcaatttg gtcccc	56
30	<210>	360	
	<211>	51	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
35	<223>	Oligonucleotide	
	<400>	360	
		cgcggatccc atatggtttc agacggccta tacaaccaac atggtaaat t	51
40	<210>	361	
	<211>	56	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
45	<223>	Oligonucleotide	
	<400>	361	
		cccgctcgag tctttcctt tcaattataa ctttagtagg ttcaatttg gtcccc	56
50	<210>	362	
	<211>	27	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
55	<223>	Oligonucleotide	
	<400>	362	
		gcggccatat gaccgtttgc acccgcg	27
	<210>	363	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>			
	<223>	Oligonucleotide		
5	<400>	363		
	gcggcctcgaa	gtcagcgggc	gttcatttct t	31
	<210>	364		
	<211>	33		
	<212>	DNA		
10	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
	<400>	364		
15	cgcgatccc	atatgaacac	cattttcaaa atc	33
	<210>	365		
	<211>	32		
	<212>	DNA		
	<213>	Artificial Sequence		
20	<220>			
	<223>	Oligonucleotide		
	<400>	365		
	cccgctcgag	ttaatttact	tttttgatgt cg	32
25	<210>	366		
	<211>	28		
	<212>	DNA		
	<213>	Artificial Sequence		
30	<220>			
	<223>	Oligonucleotide		
	<400>	366		
	gcggccatat	ggattcgccc	aaggtcgg	28
35	<210>	367		
	<211>	31		
	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
40	<400>	367		
	gcggcctcgaa	gctacacttc	ccccgaagt g	31
	<210>	368		
	<211>	31		
	<212>	DNA		
45	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
	<400>	368		
50	cgcgatccc	atatgatagt	tgaccaaagc c	31
	<210>	369		
	<211>	30		
	<212>	DNA		
	<213>	Artificial Sequence		
55	<220>			
	<223>	Oligonucleotide		

	<400> 369 cccgctcgag ttattttcc gatTTTCGG	30
5	<210> 370 <211> 28 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide	
	<400> 370 gcggccatat gcttgaactg aacggact	28
15	<210> 371 <211> 30 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
20	<400> 371 gcggcctcgA gtcagcggaa gcggacgatt	30
25	<210> 372 <211> 34 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
30	<400> 372 cgcggatccc atatgtccaa actcaaaacc atcg	34
	<210> 373 <211> 29 <212> DNA <213> Artificial Sequence	
35	<220> <223> Oligonucleotide	
	<400> 373 cccgctcgag gcttccaATC agTTTgacc	29
40	<210> 374 <211> 32 <212> DNA <213> Artificial Sequence	
45	<220> <223> Oligonucleotide	
	<400> 374 gcggccatat gagcgcaATC gttgatATT tc	32
50	<210> 375 <211> 34 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
55	<400> 375 gcggcctcgA gttatTTGCC cagTTggtag aatg	34

5	<210> 376 <211> 32 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide	
15	<400> 376 gcggccatat ggtgatacat ccgcactact tc	32
20	<210> 377 <211> 32 <212> DNA <213> Artificial Sequence	
25	<220> <223> Oligonucleotide	
30	<400> 377 gcggcctcga gtcaaaatcg agtttacac ca	32
35	<210> 378 <211> 31 <212> DNA <213> Artificial Sequence	
40	<220> <223> Oligonucleotide	
45	<400> 378 gcggccatat gaccatctat ttcaaaaacg g	31
50	<210> 379 <211> 34 <212> DNA <213> Artificial Sequence	
55	<220> <223> Oligonucleotide	
	<400> 379 gcggcctcga gtcagccat gtttagcgat catt	34
	<210> 380 <211> 31 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
	<400> 380 cgcggatccc atatgagcag cggagggggt g	31
	<210> 381 <211> 27 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
	<400> 381 ccccgctcag ttgcttggcg gcaaggc	27
	<210> 382 <211> 31	

<212>	DNA	
<213>	Artificial Sequence	
5		
<220>		
<223>	Oligonucleotide	
<400>	382	
	cgcgatccc atatggtcgc cgccgacatc g	31
10		
<210>	383	
<211>	27	
<212>	DNA	
<213>	Artificial Sequence	
15		
<220>		
<223>	Oligonucleotide	
<400>	383	
	cccgctcgag ttgcttggcg gcaaggc	27
20		
<210>	384	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Oligonucleotide	
25		
<400>	384	
	cgcgatccc atatggcgg ttcggaaggc g	31
<210>	385	
<211>	33	
<212>	DNA	
<213>	Artificial Sequence	
30		
<220>		
<223>	Oligonucleotide	
<400>	385	
	cccgctcgag ttgaacactg atgtcttttc cga	33
35		
<210>	386	
<211>	35	
<212>	DNA	
<213>	Artificial Sequence	
40		
<220>		
<223>	Oligonucleotide	
<400>	386	
	cgcgatccg ctagcaaact gtcgttggtg ttaac	35
45		
<210>	387	
<211>	26	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Oligonucleotide	
50		
<400>	387	
	cccgctcgag ttgacccgct ccacgg	26
55		
<210>	388	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	

	<220>		
	<223>	Oligonucleotide	
5	<400>	388	
	gccccat ggcggacttg gcgcaagacc c		31
	<210>	389	
	<211>	39	
	<212>	DNA	
10	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	389	
15	gcccctcga gatctcctaa acctgtttta acaatgccg		39
	<210>	390	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
20	<220>		
	<223>	Oligonucleotide	
	<400>	390	
	gccccat ggcggacttg gcgcaagacc c		31
25	<210>	391	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
30	<220>		
	<223>	Oligonucleotide	
	<400>	391	
	gcggccat gctccatgct gttgccccag c		31
35	<210>	392	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
40	<400>	392	
	gccccat ggcggacttg gcgcaagacc c		31
	<210>	393	
	<211>	31	
	<212>	DNA	
45	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	393	
50	gcggccat gaaaatcccc gctaaccgca g		31
	<210>	394	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
55	<220>		
	<223>	Oligonucleotide	

	<400> 394 cgcggatccc atatgagcag cggagggggt g	31
5	<210> 395 <211> 27 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide	
	<400> 395 cccgctcgag ttgcttggcg gcaaggc	27
15	<210> 396 <211> 31 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
20	<400> 396 cgcggatccc atatggtcgc cgccgacatc g	31
	<210> 397 <211> 27 <212> DNA <213> Artificial Sequence	
25	<220> <223> Oligonucleotide	
	<400> 397 cccgctcgag ttgcttggcg gcaaggc	27
30	<210> 398 <211> 33 <212> DNA <213> Artificial Sequence	
35	<220> <223> Oligonucleotide	
	<400> 398 cgcggatccc atatggacgg ttttgtgcct gtt	33
40	<210> 399 <211> 29 <212> DNA <213> Artificial Sequence	
45	<220> <223> Oligonucleotide	
	<400> 399 cccgctcgag cttacggatc aaattgacg	29
50	<210> 400 <211> 33 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
55	<400> 400 cgcggatccc atatggcag ccaatctgaa gaa	33

	<210> 401	
5	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 401	
10	cccgctcgag ctcagtttt gccgtcaa	28
	<210> 402	
	<211> 33	
	<212> DNA	
	<213> Artificial Sequence	
15	<220>	
	<223> Oligonucleotide	
	<400> 402	
	cgcggatccg ctagctactc atccattgtc cgc	33
20	<210> 403	
	<211> 29	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
25	<223> Oligonucleotide	
	<400> 403	
	cccgctcgag ccagttgtag cctattttgc	29
30	<210> 404	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
35	<400> 404	
	cgcggatccg ctagcatgct cttcacacac ac	32
	<210> 405	
	<211> 30	
	<212> DNA	
40	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 405	
	cccgctcgag ttaccaggttg tagcctattt	30
45	<210> 406	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
50	<220>	
	<223> Oligonucleotide	
	<400> 406	
	gccggccatat ggcacaaacg gaaggtttg aa	32
55	<210> 407	
	<211> 36	
	<212> DNA	

	<213> Artificial Sequence	
5	<220>	
	<223> oligonucleotide	
	<400> 407	
	gccgcctcga gaaaactgta acgcagggtt gccgtc	36
10	<210> 408	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
15	<400> 408	
	gcggccatat ggaagaaaca ccgcgcgaac cg	32
	<210> 409	
	<211> 32	
20	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 409	
25	gcggcctcga ggaacgtttt attaaactcg ac	32
	<210> 410	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> Oligonucleotide	
	<400> 410	
	gcggccatat gagaaaacctg accgataccc ta	32
35	<210> 411	
	<211> 33	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
40	<223> Oligonucleotide	
	<400> 411	
	gcggcctcga gtcaacgcca ctgccagcgg ttg	33
	<210> 412	
	<211> 48	
45	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
50	<400> 412	
	cgcggatccc atatgaagaa gaacatattg gaattttggg tcggactg	48
	<210> 413	
	<211> 38	
55	<212> DNA	
	<213> Artificial Sequence	
	<220>	

	<223> oligonucleotide	
5	<400> 413 cccgctcgag ttattcggcg gcttttccg cattgccg	38
	<210> 414	
	<211> 103	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> oligonucleotide	
	<400> 414 gggaattcca tatgaaaaag acagctatcg cgattgcagt ggcactggct ggtttcgcta ccgtacgcga ggccgcttagc gctttccgcg tggccggcgg tgc	60 103
15	<210> 415	
	<211> 38	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223> oligonucleotide	
	<400> 415 cccgctcgag ttattcggcg gcttttccg cattgccg	38
25	<210> 416	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> oligonucleotide	
	<400> 416 catgccatgg cttccgcgt ggccggcgt gc	32
35	<210> 417	
	<211> 38	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
40	<400> 417 cccgctcgag ttattcggcg gcttttccg cattgccg	38
	<210> 418	
	<211> 31	
	<212> DNA	
45	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
	<400> 418 cgcgatccc atatgttgc cgaaacccgc c	31
50	<210> 419	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
55	<220>	
	<223> oligonucleotide	

	<400> 419 cccgctcgag aggttgtgtt ccaggttg	28
5	<210> 420 <211> 31 <212> DNA <213> Artificial Sequence	
10	<220> <223> oligonucleotide	
	<400> 420 cgcggatccc atatgaaaaa aaccgcctat g	31
15	<210> 421 <211> 28 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
20	<400> 421 cccgctcgag ttaaggttgt gttccagg	28
	<210> 422 <211> 33 <212> DNA <213> Artificial Sequence	
25	<220> <223> oligonucleotide	
	<400> 422 cgcggatccc atatgaaaaa atacctattc cgc	33
30	<210> 423 <211> 27 <212> DNA <213> Artificial Sequence	
35	<220> <223> oligonucleotide	
	<400> 423 cccgctcgag ttacgggcgg tattcgg	27
40	<210> 424 <211> 34 <212> DNA <213> Artificial Sequence	
45	<220> <223> oligonucleotide	
	<400> 424 cgcggatccc atatgcaaag caagagcatc caaa	34
50	<210> 425 <211> 27 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
55	<400> 425 cccgctcgag ttacgggcgg tattcgg	27

5	<210> 426 <211> 86 <212> DNA <213> Artificial Sequence	
10	<220> <223> oligonucleotide  <400> 426 gggaattcca tatgaaaacc ttcttcaaaa cccttccgc cgccgcgcta gcgctcatcc tcgcccgcctg ccaaagcaag agcatac 60 86	
15	<210> 427 <211> 38 <212> DNA <213> Artificial Sequence	
20	<220> <223> oligonucleotide  <400> 427 cccgctcgag ttacgggcgg tattcgggct tcataccg 38	
25	<210> 428 <211> 33 <212> DNA <213> Artificial Sequence	
30	<220> <223> oligonucleotide  <400> 428 cgcggatccg tcgactgtgg gggcggcggt ggc 33	
35	<210> 429 <211> 31 <212> DNA <213> Artificial Sequence	
40	<220> <223> oligonucleotide  <400> 429 cccgctcgag tcaatcctgc tctttttgc c 31	
45	<210> 430 <211> 30 <212> DNA <213> Artificial Sequence	
50	<220> <223> oligonucleotide  <400> 430 gcggccatat gaagaaaaca ttgacactgc 30	
55	<210> 431 <211> 32 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide  <400> 431 gcggcctcgaa gttaatggtg cgaatgaccg at 32	
	<210> 432 <211> 46	

	<212>	DNA		
	<213>	Artificial Sequence		
5	<220>			
	<223>	Oligonucleotide		
	<400>	432		
	ggggacaagt ttgtacaaaa aaggcaggctt gcggcaagga tgccgg		46	
10	<210>	433		
	<211>	47		
	<212>	DNA		
	<213>	Artificial Sequence		
15	<220>			
	<223>	Oligonucleotide		
	<400>	433		
	ggggaccact ttgtacaaga aagctgggtc taaagcaaca atgccgg		47	
20	<210>	434		
	<211>	30		
	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
25	<400>	434		
	cgcggatccc atatgaaaca caccgtatcc		30	
	<210>	435		
	<211>	26		
	<212>	DNA		
30	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
	<400>	435		
	cccgctcgag ttatctcgta cgcgcc		26	
35	<210>	436		
	<211>	30		
	<212>	DNA		
	<213>	Artificial Sequence		
40	<220>			
	<223>	Oligonucleotide		
	<400>	436		
	cgcggatccc atatgagccc cgcgccgatt		30	
45	<210>	437		
	<211>	28		
	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
50	<223>	Oligonucleotide		
	<400>	437		
	cccgctcgag ttttgtgcg gtcaggcg		28	
55	<210>	438		
	<211>	62		
	<212>	DNA		
	<213>	Artificial Sequence		

	<220>			
	<223>	Oligonucleotide		
5	<400>	438		
	ggggacaagt ttgtacaaaa aagcaggctt gttcggttgg gggatttaaa ccaaaccaaa		60	
	tc		62	
	<210>	439		
	<211>	30		
	<212>	DNA		
10	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
	<400>	439		
15	cgcggatccc atatggcgga tgcgccccg		30	
	<210>	440		
	<211>	26		
	<212>	DNA		
20	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
	<400>	440		
	cccgctcgag aaaccgccaa tccgcc		26	
25	<210>	441		
	<211>	61		
	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
30	<223>	Oligonucleotide		
	<400>	441		
	ggggaccact ttgtacaaga aagctgggtt cattttgttt ttccttcttc tcgaggccat		60	
	t		61	
35	<210>	442		
	<211>	30		
	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
40	<223>	Oligonucleotide		
	<400>	442		
	cgcggatccc atatgaaacc caaaccgcac		30	
	<210>	443		
	<211>	27		
45	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
	<223>	Oligonucleotide		
50	<400>	443		
	cccgctcgag tcagcggtgg acgttagt		27	
	<210>	444		
	<211>	33		
	<212>	DNA		
55	<213>	Artificial Sequence		
	<220>			

	<223> Oligonucleotide	
5	<400> 444 gggaattcca tatgaaaaaa atcatcttcg ccg	33
	<210> 445	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> Oligonucleotide	
	<400> 445 cccgctcgag ttattgtttg gctgcctcga t	31
15	<210> 446	
	<211> 33	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223> Oligonucleotide	
	<400> 446 gggaattcca tatggccacc tacaaagtgg acg	33
25	<210> 447	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
30	<400> 447 cggggatcct tgtttgctg cctcgatttg	30
	<210> 448	
	<211> 34	
	<212> DNA	
35	<213> Artificial sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 448 cgcgatccc atatgcaaga acaatcgac aaag	34
40	<210> 449	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
45	<220>	
	<223> Oligonucleotide	
	<400> 449 cccgctcgag tttttcggc aaattggctt	30
50	<210> 450	
	<211> 45	
	<212> DNA	
	<213> Artificial Sequence	
55	<220>	
	<223> Oligonucleotide	
	<400> 450	

	ggggacaagt ttgtacaaaa aagcaggctg ccgatgccgt tgcgg	45
5	<210> 451	
	<211> 47	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
10	<400> 451	
	ggggaccact ttgtacaaga aagctgggtt cagggtcgtt tgttgcg	47
	<210> 452	
	<211> 30	
	<212> DNA	
15	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 452	
20	cgcggatccc atatgaaaca ctttccatcc	30
	<210> 453	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
25	<220>	
	<223> Oligonucleotide	
	<400> 453	
	cccgctcgag ttaccactcg taattgac	28
30	<210> 454	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
35	<223> Oligonucleotide	
	<400> 454	
	cgcggatccc atatggccac aagcgacgac	30
	<210> 455	
	<211> 28	
	<212> DNA	
40	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
45	<400> 455	
	cccgctcgag ttaccactcg taattgac	28
	<210> 456	
	<211> 28	
	<212> DNA	
50	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 456	
55	cgcggatccc atatggccac aaacgacg	28
	<210> 457	

	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
5	<220>		
	<223>	Oligonucleotide	
	<400>	457	
		cccgctcgag acccacgttg taaggttg	28
10	<210>	458	
	<211>	32	
	<212>	DNA	
	<213>	Artificial Sequence	
15	<220>		
	<223>	Oligonucleotide	
	<400>	458	
		cgcggatccc atatggccac aagcgacgac ga	32
20	<210>	459	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
25	<400>	459	
		cccgctcgag acccacgttg taaggttg	28
	<210>	460	
	<211>	33	
	<212>	DNA	
30	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	460	
35		cgcggatccc atatgatgaa acactttcca tcc	33
	<210>	461	
	<211>	29	
	<212>	DNA	
	<213>	Artificial Sequence	
40	<220>		
	<223>	Oligonucleotide	
	<400>	461	
		cccgctcgag ttaacccacg ttgtaaggt	29
45	<210>	462	
	<211>	33	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
50	<223>	Oligonucleotide	
	<400>	462	
		cgcggatccc atatgatgaa acactttcca tcc	33
55	<210>	463	
	<211>	29	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	oligonucleotide	
5	<400>	463	
	cccgctcgag	ttaacccacg	29
	<210>	464	
	<211>	28	
	<212>	DNA	
10	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
15	<400>	464	
	cgcggatccc	atatggccac	28
	<210>	465	
	<211>	29	
	<212>	DNA	
	<213>	Artificial Sequence	
20	<220>		
	<223>	oligonucleotide	
	<400>	465	
	cccgctcgag	gtctgacact	29
	<210>	465	
	<211>	29	
	<212>	DNA	
25	<213>	Artificial Sequence	
	<220>		
30	<223>	oligonucleotide	
	<400>	466	
	cgcggatccc	atatgatgaa	33
	<210>	466	
	<211>	33	
	<212>	DNA	
35	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
40	<400>	467	
	cccgctcgag	ttatgcttg	29
	<210>	467	
	<211>	29	
	<212>	DNA	
45	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
50	<400>	468	
	cgcggatccc	atatggccac	30
	<210>	468	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
55	<220>		
	<223>	oligonucleotide	

	<400> 469 cgcggatccc cactcgtaat tgacgcc	27
5	<210> 470 <211> 30 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide	
	<400> 470 cgcggatccc atatggccac aagcgacgac	30
15	<210> 471 <211> 27 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
20	<400> 471 cgcggatccc cactcgtaat tgacgcc	27
	<210> 472 <211> 30 <212> DNA <213> Artificial Sequence	
25	<220> <223> oligonucleotide	
	<400> 472 cgcggatccc atatggccac aaacgacgac	30
30	<210> 473 <211> 27 <212> DNA <213> Artificial Sequence	
35	<220> <223> Oligonucleotide	
	<400> 473 cgcggatcca cccacgttgt aaggttg	27
40	<210> 474 <211> 33 <212> DNA <213> Artificial Sequence	
45	<220> <223> Oligonucleotide	
	<400> 474 cgcggatccc atatgtgaa acactttcca tcc	33
50	<210> 475 <211> 27 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
55	<400> 475 cgcggatcca cccacgttgt aaggttg	27

	<210>	476	
	<211>	25	
	<212>	DNA	
5	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	476	
		cgcgatccg gaggggtgg tgtcg	25
10	<210>	477	
	<211>	27	
	<212>	DNA	
15	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	477	
		cccgctcgag ttgcttggcg gcaaggc	27
20	<210>	478	
	<211>	25	
	<212>	DNA	
25	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	478	
		cgcggatccg gcggaggcgg cactt	25
30	<210>	479	
	<211>	26	
	<212>	DNA	
35	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	479	
		cccgctcgag gaaccggtag cctacg	26
40	<210>	480	
	<211>	41	
	<212>	DNA	
45	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	480	
		cgcggatccg gtgggtgg ttcagatttgc aaacgatt c	41
50	<210>	481	
	<211>	29	
	<212>	DNA	
55	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	481	
		cccgctcgag cgtatcatat ttcacgtgc	29
	<210>	482	
	<211>	25	

	<212> DNA	
	<213> Artificial Sequence	
5	<220>	
	<223> Oligonucleotide	
	<400> 482	
	cgcgatccg gaggggtgg tgtcg	25
10	<210> 483	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
15	<220>	
	<223> Oligonucleotide	
	<400> 483	
	cccgctcgag ttattgcttg gcggcaag	28
20	<210> 484	
	<211> 25	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
25	<400> 484	
	cgcgatccg gcgaggcgg cactt	25
	<210> 485	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> Oligonucleotide	
	<400> 485	
	cccgctcgag tcagaaccgg tagcctac	28
35	<210> 486	
	<211> 41	
	<212> DNA	
	<213> Artificial Sequence	
40	<220>	
	<223> Oligonucleotide	
	<400> 486	
	cgcgatccg gtgggtgg ttcagatttgc aaacgatt c	41
45	<210> 487	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
50	<400> 487	
	cccgctcgag ttacgtatca tatttcacgt gc	32
	<210> 488	
	<211> 31	
	<212> DNA	
55	<213> Artificial Sequence	

	<220>		
	<223>	oligonucleotide	
5	<400>	488	
	cgcgatccc atatggccac aagcgacgac g		31
	<210>	489	
	<211>	28	
	<212>	DNA	
10	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	489	
15	cccgctcgag ccactcgtaa ttgacgcc		28
	<210>	490	
	<211>	30	
	<212>	DNA	
	<213>	Artificial Sequence	
20	<220>		
	<223>	oligonucleotide	
	<400>	490	
	cgcgatccc atatggccac aaacgacgac		30
25	<210>	491	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
30	<223>	oligonucleotide	
	<400>	491	
	cccgctcgag tgctttggcg gcaaagggtt		28
	<210>	492	
	<211>	30	
35	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
40	<400>	492	
	cgcgatccc atatggccac aaacgacgac		30
	<210>	493	
	<211>	37	
45	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	493	
50	cccgctcgag ttttagcaata ttatcttgt tcgtagc		37
	<210>	494	
	<211>	32	
	<212>	DNA	
	<213>	Artificial Sequence	
55	<220>		
	<223>	oligonucleotide	

	<400> 494		
	cgcggatccc atatgaaagc aaaccgtgcc ga		32
5	<210> 495		
	<211> 28		
	<212> DNA		
	<213> Artificial Sequence		
10	<220>		
	<223> Oligonucleotide		
	<400> 495		
	cccgctcgag ccactcgtaa ttgacgccc		28
15	<210> 496		
	<211> 61		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
	<223> Oligonucleotide		
20	<400> 496		
	ggggacaagt ttgtacaaaa aaggcaggctg cagccacaaa cgacgacgat gttaaaaaag		60
	C		61
	<210> 497		
	<211> 61		
	<212> DNA		
25	<213> Artificial Sequence		
	<220>		
	<223> Oligonucleotide		
	<400> 497		
30	ggggaccact ttgtacaaga aagctgggtt taccactcgt aattgacgcc gacatggtag		60
	g		61
	<210> 498		
	<211> 31		
	<212> DNA		
35	<213> Artificial Sequence		
	<220>		
	<223> Oligonucleotide		
	<400> 498		
40	gcggccatat ggcagcaaaa gacgtacagt t		31
	<210> 499		
	<211> 33		
	<212> DNA		
	<213> Artificial Sequence		
45	<220>		
	<223> Oligonucleotide		
	<400> 499		
	gcggccatcgaa gttacatcat gccgccccata cca		33
50	<210> 500		
	<211> 31		
	<212> DNA		
	<213> Artificial Sequence		
55	<220>		
	<223> Oligonucleotide		
	<400> 500		

	<b>cgcggatccg ctagcttagg cggcgccgga g</b>	<b>31</b>
5	<210> 501	
	<211> 26	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
10	<400> 501	
	<b>cccgcgtcgag gaaccggtag cctacg</b>	<b>26</b>
	<210> 502	
	<211> 29	
	<212> DNA	
15	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
	<400> 502	
20	<b>ccccctagcta gcacttctgc gcccgactt</b>	<b>29</b>
	<210> 503	
	<211> 26	
	<212> DNA	
	<213> Artificial Sequence	
25	<220>	
	<223> Oligonucleotide	
	<400> 503	
	<b>cccgcgtcgag gaaccggtag cctacg</b>	<b>26</b>
30	<210> 504	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
35	<223> Oligonucleotide	
	<400> 504	
	<b>cgcggatccg ctagcttagg cggcgccgga g</b>	<b>31</b>
40	<210> 505	
	<211> 26	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
45	<400> 505	
	<b>cccgcgtcgag gaaccggtag cctacg</b>	<b>26</b>
	<210> 506	
	<211> 32	
	<212> DNA	
50	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 506	
55	<b>cgcggatccg ctagcacttc tgcgcccgac tt</b>	<b>32</b>
	<210> 507	

<211>	26	
<212>	DNA	
<213>	Artificial Sequence	
5		
<220>		
<223>	Oligonucleotide	
<400>	507	
	<code>cccgctcgag gaaccggtag cctacg</code>	26
10		
<210>	508	
<211>	50	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
15	<223>	Oligonucleotide
<400>	508	
	<code>cgcggatccg cttagccgaac gaccccaacc ttcccataaa aaactttcaa</code>	50
20		
<210>	509	
<211>	35	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
25	<223>	Oligonucleotide
<400>	509	
	<code>cccgctcgag tcagaaccga cgtgccaagc cgttc</code>	35
30		
<210>	510	
<211>	32	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
35	<223>	Oligonucleotide
<400>	510	
	<code>gccgcatat gcccccactg gaagaacgga cg</code>	32
40		
<210>	511	
<211>	35	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
45	<223>	Oligonucleotide
<400>	511	
	<code>gccgcctcgaa gtaataaacc ttctatgggc agcag</code>	35
50		
<210>	512	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
55	<223>	Oligonucleotide
<400>	512	
	<code>cgcggatccc atatgtccgt ccacgcatcc g</code>	31
<210>	513	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	

	<220>											
	<223>	Oligonucleotide										
5	<400>	513										
	cccgctcgag	tttgaatttg taggtgtatt g	31									
	<210>	514										
	<211>	29										
	<212>	DNA										
10	<213>	Artificial Sequence										
	<220>											
	<223>	Oligonucleotide										
	<400>	514										
15	cgcggatccc	atatgacccc ttccgcact	29									
	<210>	515										
	<211>	32										
	<212>	DNA										
	<213>	Artificial Sequence										
20	<220>											
	<223>	Oligonucleotide										
	<400>	515										
	cccgctcgag	ttatattgaat ttgttaggtgt at	32									
25	<210>	516										
	<211>	33										
	<212>	DNA										
	<213>	Artificial Sequence										
30	<220>											
	<223>	Oligonucleotide										
	<400>	516										
	cgcggatccc	atatgaaaac caattcagaa gaa	33									
35	<210>	517										
	<211>	28										
	<212>	DNA										
	<213>	Artificial Sequence										
	<220>											
	<223>	Oligonucleotide										
40	<400>	517										
	cccgctcgag	tccacagaga ttgtttcc	28									
	<210>	518										
	<211>	17										
45	<212>	DNA										
	<213>	Artificial Sequence										
	<220>											
	<223>	Oligonucleotide										
50	<400>	518										
	gatccccgaa	gggcggg	17									
	<210>	519										
	<211>	29										
	<212>	DNA										
55		<213>	Artificial Sequence			<220>				<223>	Oligonucleotide	
	<213>	Artificial Sequence										
	<220>											
	<223>	Oligonucleotide										

	<400> 519 gccccaaagctt tcagaagaag acttcacgc	29
5	<210> 520 <211> 36 <212> DNA <213> Artificial Sequence	
10	<220> <223> oligonucleotide	
	<400> 520 cgcggatccc atatgcaaac ccataaatac gctatt	36
15	<210> 521 <211> 29 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
20	<400> 521 gccccaaagctt gaagaagact tcacgccag	29
25	<210> 522 <211> 35 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
30	<400> 522 cgcggatccc atatggtctt tttcgacaat accga	35
	<210> 523 <211> 10 <212> DNA <213> Artificial Sequence	
35	<220> <223> oligonucleotide	
	<400> 523 gccccaaagctt	10
40	<210> 524 <211> 36 <212> DNA <213> Artificial Sequence	
45	<220> <223> oligonucleotide	
	<400> 524 cgcggatccc atatgaataa aactttaaaa aggccgg	36
50	<210> 525 <211> 29 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
55	<400> 525 gccccaaagctt tcagaagaag acttcacgc	29

5	<210> 526 <211> 35 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide	
15	<400> 526 cgcgaatccc atatgttcga tcttgattct gtcga	35
20	<210> 527 <211> 28 <212> DNA <213> Artificial Sequence	
25	<220> <223> Oligonucleotide	
30	<400> 527 cccgctcgag tcgcacaggc tgttggcg	28
35	<210> 528 <211> 32 <212> DNA <213> Artificial Sequence	
40	<220> <223> Oligonucleotide	
45	<400> 528 cgcgaatccc atatgttggg cgaggcgcc ag	32
50	<210> 529 <211> 28 <212> DNA <213> Artificial Sequence	
55	<220> <223> Oligonucleotide	
	<400> 529 cccgctcgag tcgcacaggc tgttggcg	28
	<210> 530 <211> 32 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
	<400> 530 cgcgaatccc atatgttggg cgaggcgcc ag	32
	<210> 531 <211> 28 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
	<400> 531 cccgctcgag tcgcacaggc tgttggcg	28
	<210> 532 <211> 33	

	<212> DNA	
	<213> Artificial Sequence	
5	<220>	
	<223> oligonucleotide	
	<400> 532	
	cgcgatccc atatggcaaa tttggaggtg cgc	33
10	<210> 533	
	<211> 27	
	<212> DNA	
	<213> Artificial Sequence	
15	<220>	
	<223> oligonucleotide	
	<400> 533	
	cccgctcgag ttccggagcgg ttgaagc	27
20	<210> 534	
	<211> 34	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
25	<400> 534	
	cgcgatccc atatgcaacg tcgtattata accc	34
	<210> 535	
	<211> 29	
	<212> DNA	
	<213> Artificial Sequence	
30	<220>	
	<223> oligonucleotide	
	<400> 535	
	cccgctcgag ttattcgag cggttgaag	29
35	<210> 536	
	<211> 42	
	<212> DNA	
	<213> Artificial Sequence	
40	<220>	
	<223> oligonucleotide	
	<400> 536	
	cgcggatccc atatggcat caaagtgcgc atcaacggct ac	42
45	<210> 537	
	<211> 35	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
50	<400> 537	
	cccgctcgag tttgagcggtt cgcaactcaa gtcccg	35
	<210> 538	
	<211> 33	
	<212> DNA	
	<213> Artificial Sequence	

	<220>		
	<223>	Oligonucleotide	
5	<400>	538	
	cgcggatccc atatgggcgg cagcgaaaaa aac		33
	<210>	539	
	<211>	28	
	<212>	DNA	
10	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	539	
	cccgctcgag gttggtgccg actttgat		28
15	<210>	540	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
20	<220>		
	<223>	Oligonucleotide	
	<400>	540	
	cgcggatccc atatgggcgg cgaaagcgat a		31
25	<210>	541	
	<211>	27	
	<212>	DNA	
	<213>	Artificial Sequence	
30	<220>		
	<223>	Oligonucleotide	
	<400>	541	
	cccgctcgag tttgccccgt ttgagcc		27
35	<210>	542	
	<211>	33	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
40	<400>	542	
	cgcggatccc atatggcaa atccgaaaat acg		33
	<210>	543	
	<211>	27	
	<212>	DNA	
45	<213>	Artificial Sequence	
	<220>		
	<223>	Oligonucleotide	
	<400>	543	
	cccgctcgag catcccgtag tgtttcg		27
50	<210>	544	
	<211>	62	
	<212>	DNA	
	<213>	Artificial Sequence	
55	<220>		
	<223>	Oligonucleotide	

	<400> 544	ggggacaagt ttgtacaaaa aagcaggctc cgacattacc gtgtacaacg gccaacaaag	60
	aa		62
5	<210> 545		
	<211> 61		
	<212> DNA		
	<213> Artificial Sequence		
10	<220>		
	<223> oligonucleotide		
	<400> 545	ggggaccact ttgtacaaga aagctgggtc ttatccata ccggcttgct caagcagccg	60
	g		61
15	<210> 546		
	<211> 61		
	<212> DNA		
	<213> Artificial Sequence		
20	<220>		
	<223> oligonucleotide		
	<400> 546	ggggacaagt ttgtacaaaa aagcaggctg atacggtgtt ttcctgtaaa acggacaaca	60
	a		61
25	<210> 547		
	<211> 60		
	<212> DNA		
	<213> Artificial Sequence		
30	<220>		
	<223> oligonucleotide		
	<400> 547	ggggaccact ttgtacaaga aagctgggtc tagaaaaat cgtcatcggtt gaaattcgcc	60
35	<210> 548		
	<211> 47		
	<212> DNA		
	<213> Artificial Sequence		
	<220>		
	<223> oligonucleotide		
40	<400> 548	ggggacaagt ttgtacaaaa aagcaggcta tgccccat cgaaacc	47
	<210> 549		
	<211> 47		
	<212> DNA		
45	<213> Artificial Sequence		
	<220>		
	<223> oligonucleotide		
	<400> 549	ggggaccact ttgtacaaga aagctgggtc tagtcttgca gtgcctc	47
50			
	<210> 550		
	<211> 41		
	<212> DNA		
	<213> Artificial Sequence		
55	<220>		
	<223> oligonucleotide		

	<400> 550	
	cgcggatccc atatggaaa tttcttatata agaggcatta g	41
5	<210> 551	
	<211> 40	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> Oligonucleotide	
	<400> 551	
	cccgctcgag gttaatttct atcaactctt tagcaataat	40
15	<210> 552	
	<211> 31	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
20	<400> 552	
	cgcggatccc atatggcctg ccaagacgac a	31
	<210> 553	
	<211> 26	
	<212> DNA	
25	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 553	
30	cccgctcgag ccgcctcctg ccgaaa	26
	<210> 554	
	<211> 34	
	<212> DNA	
	<213> Artificial Sequence	
35	<220>	
	<223> oligonucleotide	
	<400> 554	
	cgcggatccc atatggcaga gatctgtttg ataa	34
40	<210> 555	
	<211> 27	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
45	<223> Oligonucleotide	
	<400> 555	
	cccgctcgag cggtttccg cccaatg	27
50	<210> 556	
	<211> 30	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
55	<400> 556	
	cgcggatccc atatgcagcc ggatacggtc	30

5	<210> 557 <211> 30 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide  <400> 557 cccgctcgag aatcacttcc aacacaaaat	30
15	<210> 558 <211> 33 <212> DNA <213> Artificial Sequence  <220> <223> oligonucleotide  <400> 558 cgcggatccc atatgtggtt gctgatgaag ggc	33
20	<210> 559 <211> 28 <212> DNA <213> Artificial Sequence  <220> <223> oligonucleotide  <400> 559 cccgctcgag gactgcttca tcttctgc	28
25	<210> 560 <211> 34 <212> DNA <213> Artificial Sequence  <220> <223> oligonucleotide  <400> 560 cgcggatccc atatgaaact gatgactgtt ttgc	34
30	<210> 561 <211> 29 <212> DNA <213> Artificial Sequence  <220> <223> oligonucleotide  <400> 561 cccgctcgag tcagactgct tcattttct	29
35	<210> 562 <211> 45 <212> DNA <213> Artificial Sequence  <220> <223> oligonucleotide  <400> 562 cgcggatccc atatgagcat taaagttagcg attaacgggtt tcggc	45
40	<210> 563 <211> 40 <212> DNA	

	<213> Artificial Sequence					
5	<220>					
	<223> Oligonucleotide					
	<400> 563					
	cccgctcgag gatTTgcct gcgaagtatt ccaaagtgcg	40				
10	<210> 564					
	<211> 32					
	<212> DNA					
	<213> Artificial Sequence					
	<220>					
	<223> oligonucleotide					
15	<400> 564					
	cgcggatccg cttagccccga tgttaaatcg gc	32				
	<210> 565					
	<211> 29					
	<212> DNA					
20	<213> Artificial Sequence					
	<220>					
	<223> Oligonucleotide					
	<400> 565					
	cggggatcca tcctgctctt ttttgccgg	29				
25	<210> 566					
	<211> 42					
	<212> DNA					
	<213> Artificial Sequence					
30	<220>					
	<223> Oligonucleotide					
	<400> 566					
	cgcggatccg gtgggtgg tcaaagcaag agcatccaaa cc	42				
35	<210> 567					
	<211> 30					
	<212> DNA					
	<213> Artificial Sequence					
40	<220>					
	<223> Oligonucleotide					
	<400> 567					
	cccaaggcttt tcgggcggta ttcccggcttc	30				
45	<210> 568					
	<211> 39					
	<212> DNA					
	<213> Artificial Sequence					
	<220>					
	<223> Oligonucleotide					
50	<400> 568					
	cgcggatccg gtgggtgg tgccacctac aaagtggac	39				
	<210> 569					
	<211> 28					
	<212> DNA					
55		<213> Artificial Sequence			<220>	
	<213> Artificial Sequence					
	<220>					

	<223> oligonucleotide	
5	<400> 569 gccccaaagctt ttgtttggct gcctcgat	28
	<210> 570	
	<211> 34	
	<212> DNA	
	<213> Artificial Sequence	
10	<220>	
	<223> oligonucleotide	
	<400> 570 cgcggatccg gtgggtgg tacaaggcgac gacg	34
15	<210> 571	
	<211> 28	
	<212> DNA	
	<213> Artificial Sequence	
20	<220>	
	<223> oligonucleotide	
	<400> 571 gccccaaagctt ccactcgtaa ttgacgccc	28
25	<210> 572	
	<211> 41	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
30	<400> 572 cgcggatccg gtgggtgg ttcagatttgc gcaaacgatt c	41
	<210> 573	
	<211> 28	
	<212> DNA	
35	<213> Artificial Sequence	
	<220>	
	<223> Oligonucleotide	
	<400> 573 cccaagcttc gtatcatatt tcacgtgc	28
40	<210> 574	
	<211> 44	
	<212> DNA	
	<213> Artificial Sequence	
45	<220>	
	<223> oligonucleotide	
	<400> 574 cccaagcttg gtgggtgg tggttcagat ttggcaaacg attc	44
50	<210> 575	
	<211> 29	
	<212> DNA	
	<213> Artificial Sequence	
55	<220>	
	<223> Oligonucleotide	
	<400> 575	

	ccccgctcgag cgtatcatat ttcacgtgc	29
5	<210> 576	
	<211> 45	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
10	<400> 576	
	cccaagcttg gtgggtgg tggtaaaggc aagagcatcc aaacc	45
	<210> 577	
	<211> 28	
	<212> DNA	
15	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
	<400> 577	
20	cccgctcgag cgggcggtat tcgggctt	28
	<210> 578	
	<211> 32	
	<212> DNA	
	<213> Artificial Sequence	
25	<220>	
	<223> oligonucleotide	
	<400> 578	
	cgcggatccg ctagccccga tgttaaatcg gc	32
30	<210> 579	
	<211> 29	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
35	<223> oligonucleotide	
	<400> 579	
	cggggatcca tcctgctctt ttttgcgg	29
40	<210> 580	
	<211> 36	
	<212> DNA	
	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
45	<400> 580	
	cgcggatccg ctagcgaca cacttatttc ggcattc	36
	<210> 581	
	<211> 30	
	<212> DNA	
50	<213> Artificial Sequence	
	<220>	
	<223> oligonucleotide	
	<400> 581	
55	cgcggatccc cagcggtagc ctaatttgat	30
	<210> 582	

	<211>	41	
	<212>	DNA	
	<213>	Artificial Sequence	
5	<220>		
	<223>	oligonucleotide	
	<400>	582	
		cgcgatccg gtgggtgg ttcagattt gcaaacgatt c	41
10	<210>	583	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
15	<223>	oligonucleotide	
	<400>	583	
		cccaagcttc gtatcatatt tcacgtgc	28
20	<210>	584	
	<211>	36	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
25	<223>	oligonucleotide	
	<400>	584	
		gcggcgtcga cggtggcga ggcactggat cctcag	36
	<210>	585	
	<211>	35	
	<212>	DNA	
30	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	585	
35		ggaggcactg gatcctcaga ttggcaaac gattc	35
	<210>	586	
	<211>	29	
	<212>	DNA	
	<213>	Artificial Sequence	
40	<220>		
	<223>	oligonucleotide	
	<400>	586	
		cccgctcgag cgtatcatat ttcacgtgc	29
45	<210>	587	
	<211>	33	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
50	<223>	oligonucleotide	
	<400>	587	
		ggaattccat atgtcagatt tggcaaacga ttc	33
55	<210>	588	
	<211>	28	
	<212>	DNA	
	<213>	Artificial Sequence	

	<220>		
	<223>	oligonucleotide	
5	<400>	588	
		cgcggatccc gtatcatatt tcacgtgc	28
	<210>	589	
	<211>	25	
	<212>	DNA	
10	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	589	
15		cggggatccg ggggcggcgg tggcg	25
	<210>	590	
	<211>	30	
	<212>	DNA	
20	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	590	
25		cccaagctta tcctgctttt ttttgcggc	30
	<210>	591	
	<211>	42	
	<212>	DNA	
30	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	591	
35		cgcggatccg gtgggtgg tcaaagcaag agcatccaaa cc	42
	<210>	592	
	<211>	28	
	<212>	DNA	
40	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	592	
45		cccaagcttc gggcggtatt cgggcttc	28
	<210>	593	
	<211>	26	
	<212>	DNA	
50	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	
	<400>	593	
55		ccccaaagctt gggggcggcg gtggcg	26
	<210>	594	
	<211>	31	
	<212>	DNA	
	<213>	Artificial Sequence	
	<220>		
	<223>	oligonucleotide	

	<400> 594 cccgctcgag atcctgctct tttttgccgg c	31
5	<210> 595 <211> 45 <212> DNA <213> Artificial Sequence	
10	<220> <223> Oligonucleotide	
	<400> 595 cccaagcttg gtgggtggtgg tggtaaaagc aagagcatcc aaacc	45
15	<210> 596 <211> 28 <212> DNA <213> Artificial Sequence	
20	<220> <223> Oligonucleotide	
	<400> 596 cccgctcgag cgggcgttat tcgggctt	28
25	<210> 597 <211> 35 <212> DNA <213> Artificial Sequence	
	<220> <223> oligonucleotide	
30	<400> 597 ggaggcactg gatccgcagc cacaaacgac gacga	35
	<210> 598 <211> 36 <212> DNA <213> Artificial Sequence	
35	<220> <223> Oligonucleotide	
	<400> 598 gcggcctcgaa gggtggcgga ggcactggat ccgcag	36
40	<210> 599 <211> 28 <212> DNA <213> Artificial Sequence	
45	<220> <223> Oligonucleotide	
	<400> 599 cccgctcgag acccagcttg taaggttg	28
50	<210> 600 <211> 35 <212> DNA <213> Artificial Sequence	
	<220> <223> Oligonucleotide	
55	<400> 600 ggaggcactg gatccgcagc cacaaacgac gacga	35

	<210>	601			
5	<211>	36			
	<212>	DNA			
	<213>	Artificial Sequence			
	<220>				
	<223>	Oligonucleotide			
10	<400>	601			
	gcggcctcga	gggtggcgga	ggcactggat	ccgcag	36
	<210>	602			
	<211>	28			
	<212>	DNA			
	<213>	Artificial Sequence			
15	<220>				
	<223>	Oligonucleotide			
	<400>	602			
	cccgctcgag	ccactcgtaa	ttgacgcc	28	
20	<210>	603			
	<211>	38			
	<212>	DNA			
	<213>	Artificial Sequence			
	<220>				
25	<223>	Oligonucleotide			
	<400>	603			
	gcggcctcga	gggatccggc	ggaggcggca	cttctgcg	38
30	<210>	604			
	<211>	26			
	<212>	DNA			
	<213>	Artificial Sequence			
	<220>				
	<223>	Oligonucleotide			
35	<400>	604			
	cccgctcgag	gaaccggtag	cctacg	26	
	<210>	605			
	<211>	35			
	<212>	DNA			
40	<213>	Artificial Sequence			
	<220>				
	<223>	Oligonucleotide			
	<400>	605			
45	ggaggcactg	gatcctcaga	tttggcaaac	gattc	35
	<210>	606			
	<211>	37			
	<212>	DNA			
	<213>	Artificial Sequence			
50	<220>				
	<223>	Oligonucleotide			
	<400>	606			
	gcggcgtcga	cggtgtggcgga	ggcactggat	cctcaga	37
55	<210>	607			
	<211>	29			

<212>	DNA	
<213>	Artificial Sequence	
5	<220>	
	<223>	oligonucleotide
	<400>	607
	cccgctcgag cgtatcatat ttcacgtgc	29
10	<210>	608
	<211>	35
	<212>	DNA
	<213>	Artificial Sequence
15	<220>	
	<223>	oligonucleotide
	<400>	608
	gcggcctcgaa gggatccgga ggggggtggtg tcgcc	35
20	<210>	609
	<211>	25
	<212>	DNA
	<213>	Artificial Sequence
	<220>	
	<223>	oligonucleotide
25	<400>	609
	cccgctcgag ttgcttggcg gcaag	25
	<210>	610
	<211>	35
	<212>	DNA
	<213>	Artificial Sequence
30	<220>	
	<223>	oligonucleotide
	<400>	610
	ggaggcactg gatccgcagc cacaaacgac gacga	35
35	<210>	611
	<211>	36
	<212>	DNA
	<213>	Artificial Sequence
40	<220>	
	<223>	oligonucleotide
	<400>	611
	gcggcctcgaa gggtggcgga ggcactggat ccgcag	36
45	<210>	612
	<211>	28
	<212>	DNA
	<213>	Artificial Sequence
50	<220>	
	<223>	oligonucleotide
	<400>	612
	cccgctcgag acccagcttg taaggttg	28
55	<210>	613
	<211>	35
	<212>	DNA
	<213>	Artificial Sequence

	<220>			
	<223>	oligonucleotide		
5	<400>	613		
	ggaggcactg	gatccgcage	cacaaacgac	
	gacga		35	
	<210>	614		
	<211>	36		
	<212>	DNA		
10	<213>	Artificial Sequence		
	<220>			
	<223>	oligonucleotide		
	<400>	614		
	gcggcctcga	gggtggcgg	ggcactggat	
15		cccgag	ccgcag	36
	<210>	615		
	<211>	28		
	<212>	DNA		
	<213>	Artificial Sequence		
20	<220>			
	<223>	oligonucleotide		
	<400>	615		
	cccgcgtcgag	ccactcgtaa	ttgacgcc	
			28	
25	<210>	616		
	<211>	35		
	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
30	<223>	oligonucleotide		
	<400>	616		
	ggaggcactg	gatcctcaga	tttggcaaac	
	gattc		35	
35	<210>	617		
	<211>	37		
	<212>	DNA		
	<213>	Artificial Sequence		
	<220>			
	<223>	oligonucleotide		
40	<400>	617		
	gcggcgtcga	cgggtggcgg	ggcactggat	
	cctcaga		37	
	<210>	618		
	<211>	29		
	<212>	DNA		
45	<213>	Artificial Sequence		
	<220>			
	<223>	oligonucleotide		
	<400>	618		
50		cccgcgtcgag	cgtatcatat	
	ttcacgtgc		29	
	<210>	619		
	<211>	488		
	<212>	PRT		
	<213>	Artificial Sequence		
55	<220>			
	<223>	MC58		

<400> 619  
 Met Phe Lys Arg Ser Val Ile Ala Met Ala Cys Ile Phe Ala Leu Ser  
 1 5 10 15  
 Ala Cys Gly Gly Gly Gly Ser Pro Asp Val Lys Ser Ala Asp  
 5 20 25 30  
 Thr Leu Ser Lys Pro Ala Ala Pro Val Val Ser Glu Lys Glu Thr Glu  
 10 35 40 45  
 Ala Lys Glu Asp Ala Pro Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro  
 50 55 60  
 Ser Ala Gln Gly Ser Gln Asp Met Ala Ala Val Ser Glu Glu Asn Thr  
 65 70 75 80  
 Gly Asn Gly Gly Ala Val Thr Ala Asp Asn Pro Lys Asn Glu Asp Glu  
 15 85 90 95  
 Val Ala Gln Asn Asp Met Pro Gln Asn Ala Ala Gly Thr Asp Ser Ser  
 100 105 110  
 Thr Pro Asn His Thr Pro Asp Pro Asn Met Leu Ala Gly Asn Met Glu  
 20 115 120 125  
 Asn Gln Ala Thr Asp Ala Gly Glu Ser Ser Gln Pro Ala Asn Gln Pro  
 130 135 140  
 Asp Met Ala Asn Ala Ala Asp Gly Met Gln Gly Asp Asp Pro Ser Ala  
 25 145 150 155 160  
 Gly Gly Gln Asn Ala Gly Asn Thr Ala Ala Gln Gly Ala Asn Gln Ala  
 165 170 175  
 Gly Asn Asn Gln Ala Ala Gly Ser Ser Asp Pro Ile Pro Ala Ser Asn  
 30 180 185 190  
 Pro Ala Pro Ala Asn Gly Gly Ser Asn Phe Gly Arg Val Asp Leu Ala  
 195 200 205  
 Asn Gly Val Leu Ile Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr His  
 35 210 215 220  
 Cys Lys Gly Asp Ser Cys Ser Gly Asn Asn Phe Leu Asp Glu Glu Val  
 225 230 235 240  
 Gln Leu Lys Ser Glu Phe Glu Lys Leu Ser Asp Ala Asp Lys Ile Ser  
 40 245 250 255  
 Asn Tyr Lys Lys Asp Gly Lys Asn Asp Lys Phe Val Gly Leu Val Ala  
 260 265 270  
 Asp Ser Val Gln Met Lys Gly Ile Asn Gln Tyr Ile Ile Phe Tyr Lys  
 45 275 280 285  
 Pro Lys Pro Thr Ser Phe Ala Arg Phe Arg Arg Ser Ala Arg Ser Arg  
 290 295 300  
 Arg Ser Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn Gln Ala Asp  
 305 310 315 320  
 Thr Leu Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly His Ser Gly  
 50 325 330 335  
 Asn Ile Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr Tyr Gly Ala  
 340 345 350  
 Glu Lys Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln Gly Glu Pro  
 55 355 360 365

Ala Lys Gly Glu Met Leu Ala Gly Ala Ala Val Tyr Asn Gly Glu Val  
 370 375 380  
 5 Leu His Phe His Thr Glu Asn Gly Arg Pro Tyr Pro Thr Arg Gly Arg  
 385 390 395 400  
 Phe Ala Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp Gly Ile Ile  
 405 410 415  
 10 Asp Ser Gly Asp Asp Leu His Met Gly Thr Gln Lys Phe Lys Ala Ala  
 420 425 430  
 Ile Asp Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn Gly Ser Gly  
 435 440 445  
 15 Asp Val Ser Gly Lys Phe Tyr Gly Pro Ala Gly Glu Glu Val Ala Gly  
 450 455 460  
 Lys Tyr Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly Phe Gly Val  
 465 470 475 480  
 20 Phe Ala Gly Lys Lys Glu Gln Asp  
 485  
 <210> 620  
 <211> 427  
 <212> PRT  
 25 <213> Artificial Sequence  
 <220>  
 <223> 2996  
 <400> 620  
 30 Met Phe Glu Arg Ser Val Ile Ala Met Ala Cys Ile Phe Ala Leu Ser  
 1 5 10 15  
 Ala Cys Gly Gly Gly Gly Ser Pro Asp Val Lys Ser Ala Asp  
 20 25 30  
 35 Thr Leu Ser Lys Pro Ala Ala Pro Val Val Ala Glu Lys Glu Thr Glu  
 35 40 45  
 Val Lys Glu Asp Ala Pro Gln Ala Gly Ser Gln Gly Gln Gly Ala Pro  
 50 55 60  
 40 Ser Thr Gln Gly Ser Gln Asp Met Ala Ala Val Ser Ala Glu Asn Thr  
 65 70 75 80  
 Gly Asn Gly Gly Ala Ala Thr Thr Asp Lys Pro Lys Asn Glu Asp Glu  
 85 90 95  
 45 Gly Pro Gln Asn Asp Met Pro Gln Asn Ser Ala Glu Ser Ala Asn Gln  
 100 105 110  
 Thr Gly Asn Asn Gln Pro Ala Asp Ser Ser Asp Ser Ala Pro Ala Ser  
 115 120 125  
 50 Asn Pro Ala Pro Ala Asn Gly Gly Ser Asn Phe Gly Arg Val Asp Leu  
 130 135 140  
 Ala Asn Gly Val Leu Ile Asp Gly Pro Ser Gln Asn Ile Thr Leu Thr  
 145 150 155 160  
 His Cys Lys Gly Asp Ser Cys Asn Gly Asp Asn Leu Leu Asp Glu Glu  
 165 170 175  
 55 Ala Pro Ser Lys Ser Glu Phe Glu Asn Leu Asn Glu Ser Glu Arg Ile  
 180 185 190

Glu Lys Tyr Lys Lys Asp Gly Lys Ser Asp Lys Phe Thr Asn Leu Val  
 195 200 205  
 5 Ala Thr Ala Val Gln Ala Asn Gly Thr Asn Lys Tyr Val Ile Ile Tyr  
 210 215 220  
 Lys Asp Lys Ser Ala Ser Ser Ser Ala Arg Phe Arg Arg Ser Ala  
 225 230 235 240  
 10 Arg Ser Arg Arg Ser Leu Pro Ala Glu Met Pro Leu Ile Pro Val Asn  
 245 250 255  
 Gln Ala Asp Thr Leu Ile Val Asp Gly Glu Ala Val Ser Leu Thr Gly  
 260 265 270  
 15 His Ser Gly Asn Ile Phe Ala Pro Glu Gly Asn Tyr Arg Tyr Leu Thr  
 275 280 285  
 Tyr Gly Ala Glu Lys Leu Pro Gly Gly Ser Tyr Ala Leu Arg Val Gln  
 290 295 300  
 20 Gly Glu Pro Ala Lys Gly Glu Met Leu Ala Gly Thr Ala Val Tyr Asn  
 305 310 315 320  
 Gly Glu Val Leu His Phe His Thr Glu Asn Gly Arg Pro Tyr Pro Thr  
 325 330 335  
 25 Arg Gly Arg Phe Ala Ala Lys Val Asp Phe Gly Ser Lys Ser Val Asp  
 340 345 350  
 Gly Ile Ile Asp Ser Gly Asp Asp Leu His Met Gly Thr Gln Lys Phe  
 355 360 365  
 Lys Ala Ala Ile Asp Gly Asn Gly Phe Lys Gly Thr Trp Thr Glu Asn  
 370 375 380  
 30 Gly Gly Asp Val Ser Gly Arg Phe Tyr Gly Pro Ala Gly Glu Glu  
 385 390 395 400  
 Val Ala Gly Lys Tyr Ser Tyr Arg Pro Thr Asp Ala Glu Lys Gly Gly  
 405 410 415  
 35 Phe Gly Val Phe Ala Gly Lys Lys Glu Gln Asp  
 420 425

40

**Claims**

1. A method for the heterologous expression of a protein of the invention, in which (a) at least one domain in the protein is deleted and, optionally, (b) no fusion partner is used.
2. The method of claim 1, in which the protein of the invention is ORF46.
3. The method of claim 2, in which ORF46 is divided into a first domain (amino acids 1-433) and a second domain (amino acids 433-608).
4. The method of claim 2, in which the protein of the invention is 564.
5. The method of claim 4, in which protein 564 is divided into domains as shown in Figure 8.
- 55 6. The method of claim 1 in which the protein of the invention is 961.
7. The method of claim 6, in which protein 961 is divided into domains as shown in Figure 12.

8. The method of claim 1, in which the protein of the invention is 502 and the domain is amino acids 28 to 167 (numbered according to the MC58 sequence).
- 5 9. The method of claim 1, in which the protein of the invention is 287.
10. A method for the heterologous expression of a protein of the invention, in which (a) a portion of the N-terminal domain of the protein is deleted.
- 10 11. The method of claim 9 or claim 10, in which protein 287 is divided into domains A B & C shown in Figure 5.
- 15 12. The method of claim 11, in which (i) domain A, (ii) domains A and B, or (iii) domains A and C are deleted.
- 15 13. The method of claim 11, wherein (i) amino acids 1-17, (ii) amino acids 1-25, (iii) amino acids 1-69, or (iv) amino acids 1-106, of domain A are deleted.
- 20 14. A method for the heterologous expression of a protein of the invention, in which (a) no fusion partner is used, and (b) the protein's native leader peptide (if present) is used.
- 25 15. The method of claim 14, in which the protein of the invention is selected from the group consisting of: 111, 149, 206, 225-1, 235, 247-1, 274, 283, 286, 292, 401, 406, 502-1, 503, 519-1, 525-1, 552, 556, 557, 570, 576-1, 580, 583, 664, 759, 907, 913, 920-1, 936-1, 953, 961, 983, 989, Orf4, Orf7-1, Orf9-1, Orf23, Orf25, Orf37, Orf38, Orf40, Orf40.1, Orf40.2, Orf72-1, Orf76-1, Orf85-2, Orf91, Orf97-1, Orf119, Orf143.1, NMB0109, NMB2050, 008, 105, 117-1, 121-1, 122-1, 128-1, 148, 216, 243, 308, 593, 652, 726, 926, 982, Orf83-1 and Orf143-1.
- 30 16. A method for the heterologous expression of a protein of the invention, in which (a) the protein's leader peptide is replaced by the leader peptide from a different protein and, optionally, (b) no fusion partner is used.
17. The method of claim 16, in which the different protein is 961, ORF4, *E.coli* OmpA, or *E.carotovora* PelB, or in which the leader peptide is MKKYLFSAA.
- 35 18. The method of claim 17, in which the different protein is *E.coli* OmpA and the protein of the invention is ORF1.
19. The method of claim 17, in which the protein of the invention is 911 and the different protein is *E.carotovora* PelB or *E.coli* OmpA.
- 40 20. The method of claim 17, in which the different protein is ORF4 and the protein of the invention is 287.
21. A method for the heterologous expression of a protein of the invention, in which (a) the protein's leader peptide is deleted and, optionally, (b) no fusion partner is used.
- 45 22. The method of claim 21, in which the protein of the invention is 919.
23. A method for the heterologous expression of a protein of the invention, in which expression of a protein of the invention is carried out at a temperature at which a toxic activity of the protein is not manifested.
- 45 24. The method of claim 23, in which protein 919 is expressed at 30°C.
25. A method for the heterologous expression of a protein of the invention, in which protein is mutated to reduce or eliminate toxic activity.
- 50 26. The method of claim 25, in which the protein of the invention is 907, 919 or 922.
27. The method of claim 26, in which 907 is mutated at Glu-117 (e.g. Glu→Gly).
- 55 28. The method of claim 26, in which 919 is mutated at Glu-255 (e.g. Glu→Gly) and/or Glu-323 (e.g. Glu→Gly).
29. The method of claim 26, in which 922 is mutated at Glu-164 (e.g. Glu→Gly), Ser-213 (e.g. Ser→Gly) and/or Asn-348 (e.g. Asn→Gly).

30. A method for the heterologous expression of a protein of the invention, in which vector pSM214 is used or vector pET-24b is used.

5       31. The method of claim 30, in which the protein of the invention is 953 and the vector is pSM214.

32. A method for the heterologous expression of a protein of the invention, in which a protein is expressed or purified such that it adopts a particular multimeric form.

10      33. The method of claim 32, in which protein 953 is expressed and/or purified in monomeric form.

15      34. The method of claim 32, in which protein 961 is expressed and/or purified in tetrameric form.

35. The method of claim 32, in which protein 287 is expressed and/or purified in dimeric form.

15      36. The method of claim 32, in which protein 919 is expressed and/or purified in monomeric form.

37. A method for the heterologous expression of a protein of the invention, in which the protein is expressed as a lipidated protein.

20      38. The method of claim 37, in which the protein of the invention is 919, 287, ORF4, 406, 576, or ORF25.

39. A method for the heterologous expression of a protein of the invention, in which (a) the protein's C-terminus region is mutated and, optionally, (b) no fusion partner is used.

25      40. The method of claim 39, wherein the mutation is a substitution, an insertion, or a deletion

41. The method of claim 40, wherein the protein of the invention is 730, ORF29 or ORF46.

42. A method for the heterologous expression of a protein of the invention, in which the protein's leader peptide is mutated.

30      43. The method of claim 42, in which the protein of the invention is 919.

44. A method for the heterologous expression of a protein, in which a poly-glycine stretch within the protein is mutated.

35      45. The method of claim 44, wherein the protein is a protein of the invention.

46. The method of claim 45, wherein the protein of the invention is 287, 741, 983 or Tbp2.

40      47. The method of claim 46, wherein (Gly)<sub>6</sub> is deleted from 287 or 983.

48. The method of claim 46, wherein (Gly)<sub>4</sub> is deleted from Tbp2 or 741

49. The method of claim 47 or claim 48, wherein the leader peptide is also deleted.

45      50. The method of any preceding claim, in which the heterologous expression is in an *E.coli* host.

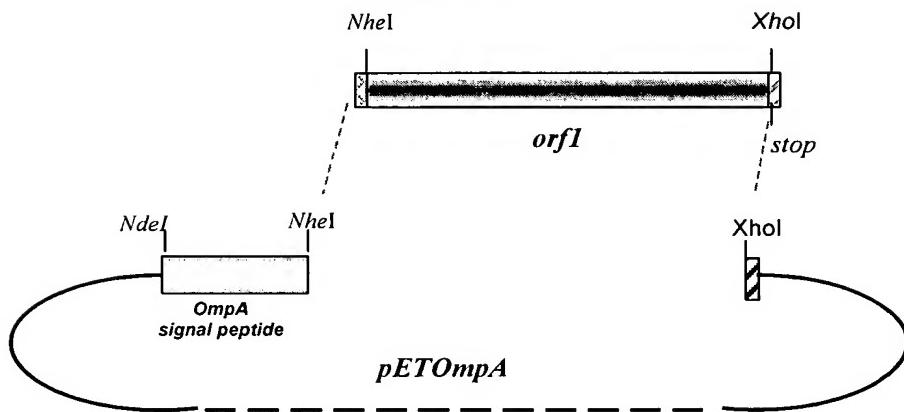
51. A protein expressed by the method of any preceding claim.

52. A heterologous protein comprising the N-terminal amino acid sequence MKKYLFSAA.

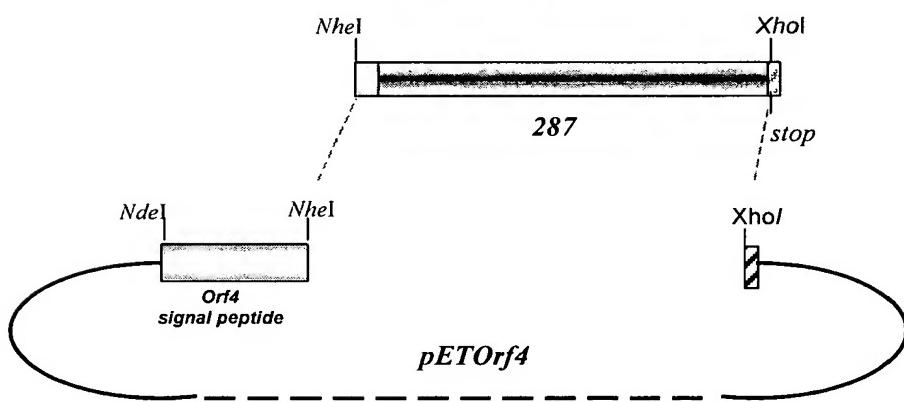
50

55

**FIGURE 1**



**FIGURE 2**



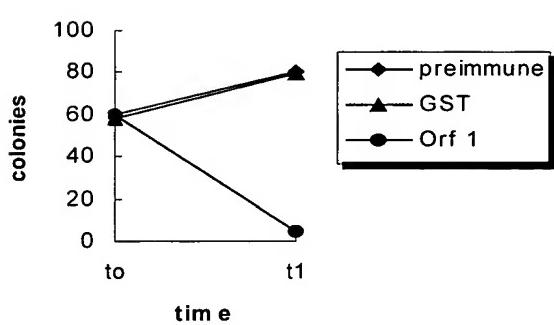
**FIGURE 3**



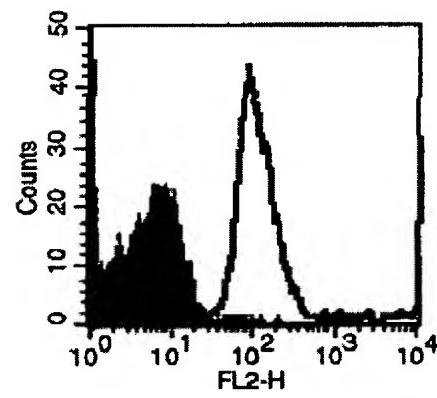
**PURIFICATION**



**WESTERN BLOT**



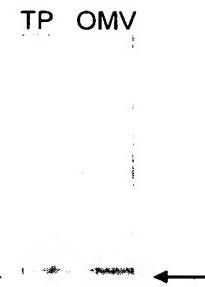
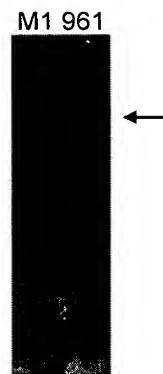
**BACTERICIDAL ASSAY**



**FACS**

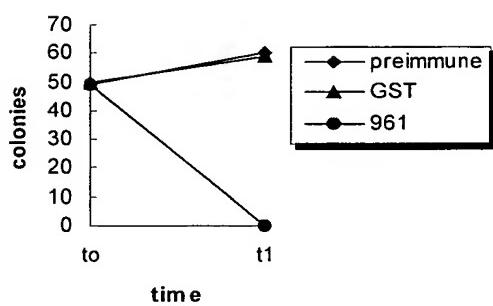
**ELISA: POSITIVE**

**FIGURE 4**

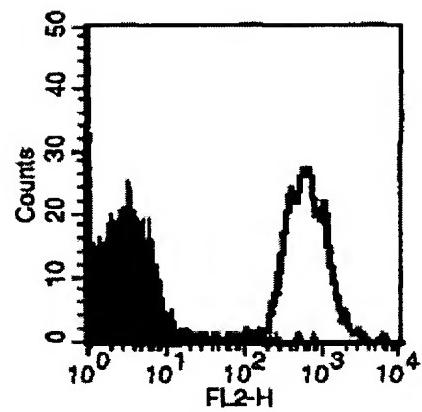


**PURIFICATION**

**WESTERN BLOT**

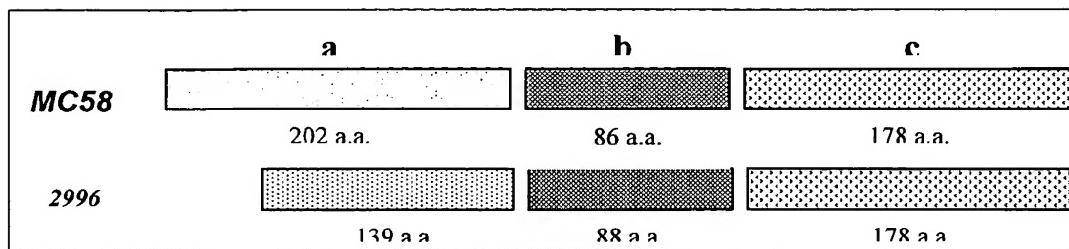
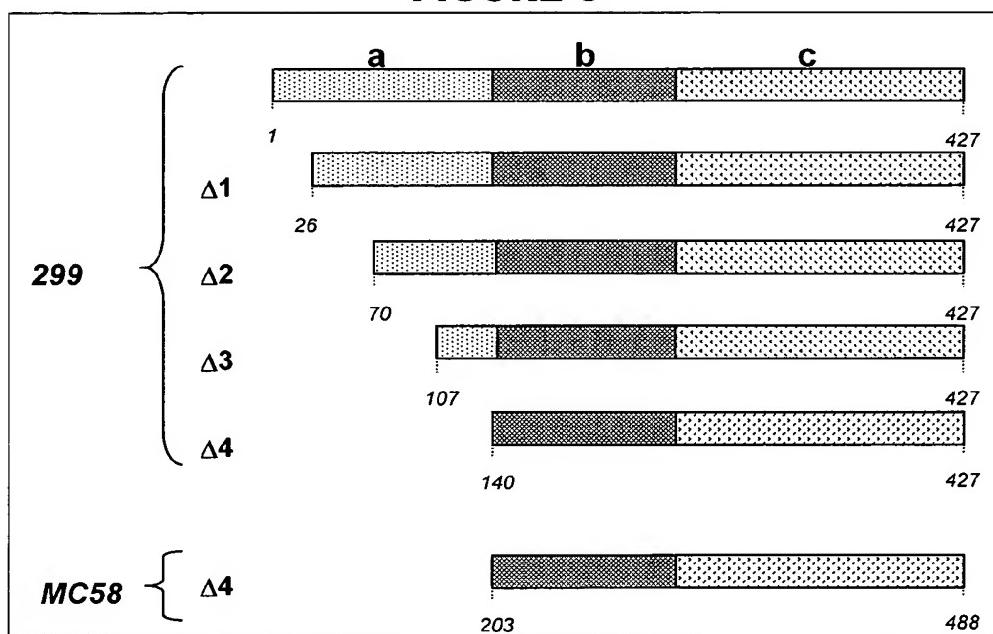
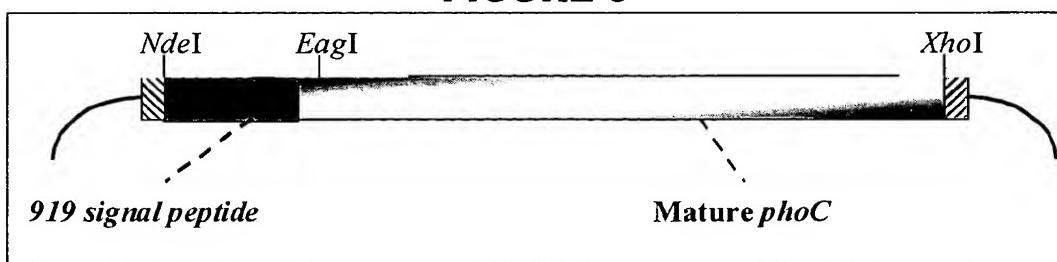


**BACTERICIDAL ASSAY**



**FACS**

**ELISA: POSITIVE**

**FIGURE 5****FIGURE 6****FIGURE 9**

**FIGURE 7**

**<A>**

MC58	1	MEKRSVIAMACIFALSACGGGGGGSPDVKSA	DTLSKPAAPVVSEKE	TEAKEDAPQAGSQG
2996	1	MEERSVIAMACIFALSACGGGGGGSPDVKSA	DTLSKPAAPVVAEKE	TEVKEDAPQAGSQG

---

**<Δ1>**

MC58	61	OGAPS	AQGSQDMAAVSEENTGN	GGAVTADNP	KNEDEVAQNDMPQNAAGTDS	STPNHTPDP
2996	61	OGAPS	TQGSQDMAAVSA	ENTGN	GGATT	DKPKNEDEGPQNDMPQN

---

**<Δ2>**

MC58	121	NMLAGNMENQATDAGESSQ	PANQPDMANAADGM	QGDDPSAGGQNAGNTAAQG	QGANOAGNNQ
2996	106	.....	.....	.....	SAESANOTGNNO

---

**<Δ3>**

MC58	181	AAGSSDPIPASNPAPANGSNEGRV	DLANGVLIDGPSQNITLTHCKGDSCSGNNELDEEV	.....	.....
2996	118	PADSSSDSAPASNPAPANGSNEGRV	DLANGVLIDGPSQNITLTHCKGDSCNGDNLLDEEA	.....	.....

---

**-A><B->**

MC58	241	QLKSEFEKLSDADKISNYKKDGKNDKFVCLVADSVQMKGINQYI	IIFYKPK..PTSEARFR	.....
2996	178	PSKSEFENLNESERIEKYKKDGKSDKFTNLVATAVQANGTNKYVIIY	KDKSASSSSARFR	.....

---

**<C->**

MC58	299	RSARSRRSLPAEMPLIPVNQADTLIVDGEAVS	LGHSGNIAFAPEGNYRYLTGAEKLPGG	.....
2996	238	RSARSRRSLPAEMPLIPVNQADTLIVDGEAVS	LGHSGNIAFAPEGNYRYLTGAEKLPGG	.....

---

MC58	359	SYALRVOGEPAKGEMLACA	AVYNGEVLFHETENGRPYPTRGRFAAKVDFGSKSVDGIIDS	.....
2996	298	SYALRVOGEPAKGEMLAGT	AVYNGEVLFHETENGRPYPTRGRFAAKVDFGSKSVDGIIDS	.....

---

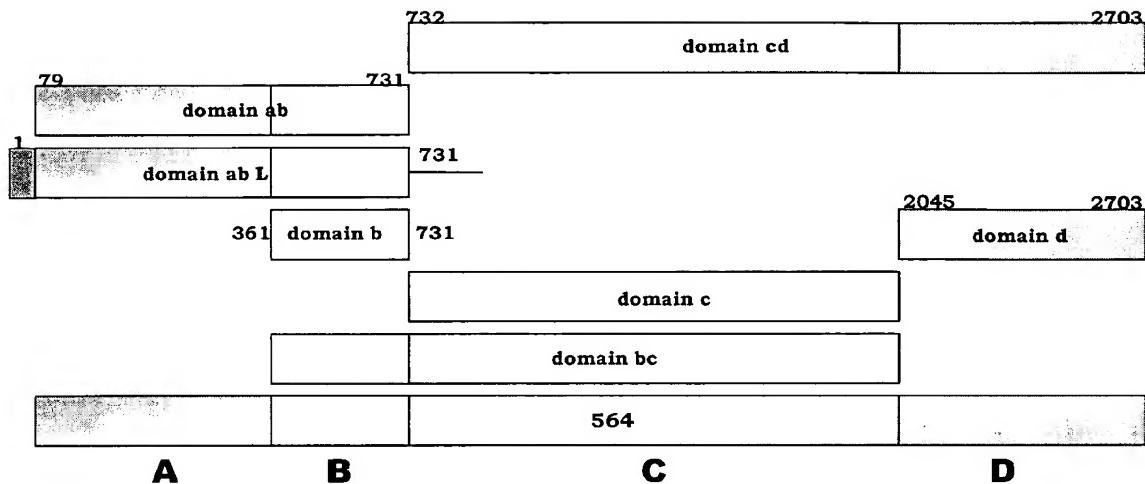
MC58	419	GDDIHMG	TQKFKAIDGNGFKGTWTENG	SGDVSGKFYGPAGEEVAGKYSYRPTDAEKGGF	.....
2996	358	GDDIHMG	TQKFKAIDGNGFKGTWTENG	GGDVSGRFYGPAGEEVAGKYSYRPTDAEKGGF	.....

---

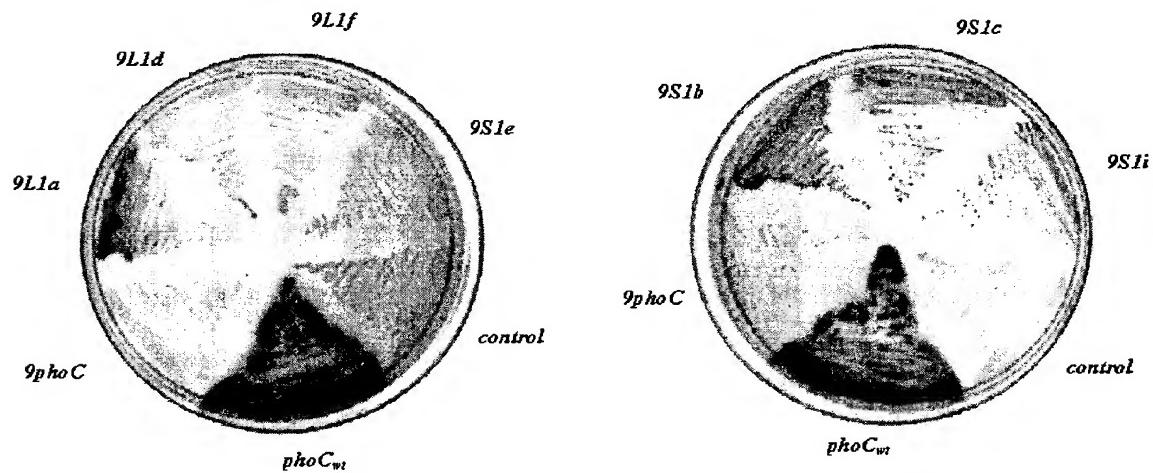
**-C->**

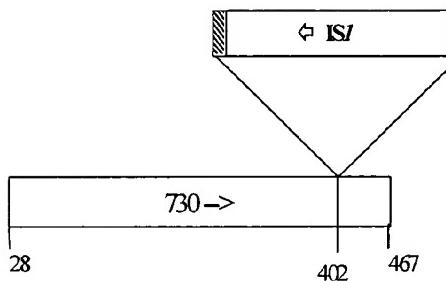
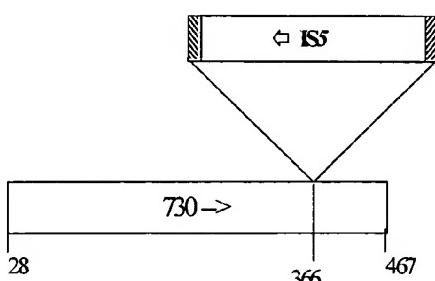
MC58	479	GVFAGKKEQD*	.....
2996	418	GVFAGKKEQD*	.....

**FIGURE 8**

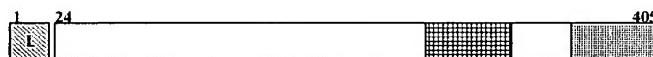


**FIGURE 10**



**FIGURE 11A****FIGURE 11B****FIGURE 12**

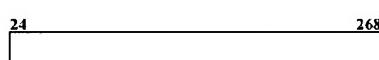
**961 (2996)**  
961 L (2996)



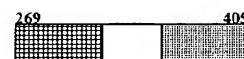
**961 (MC58)**  
961 L (MC58)



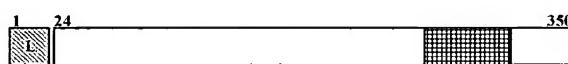
**961a (2996=MC58)**



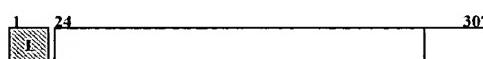
**961b (2996)**



**961c (2996)**  
961c-L (2996)



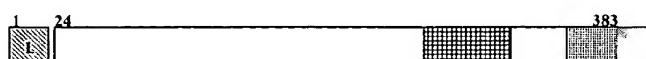
**961c (MC58)**  
961c-L (MC58)



**961d (2996)**



**961-Δ1 (2996)**  
961Δ1-L

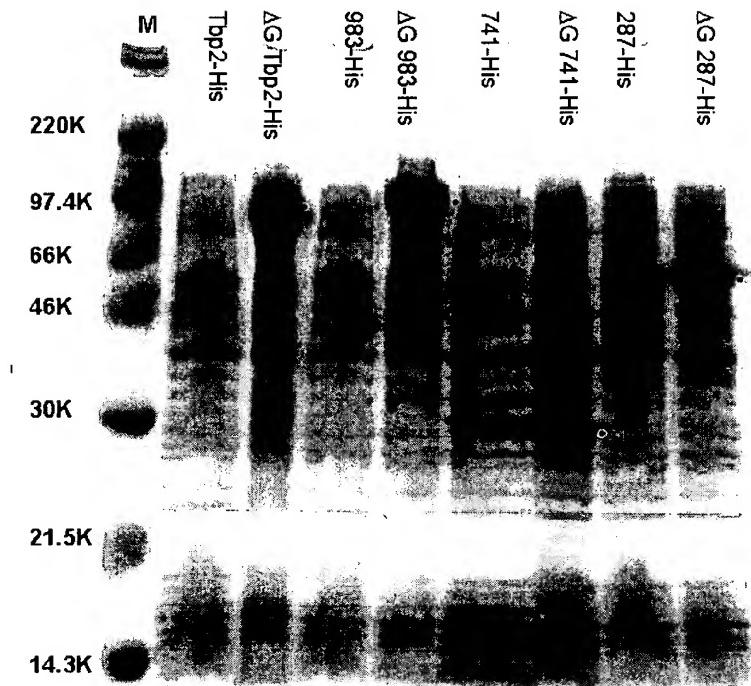
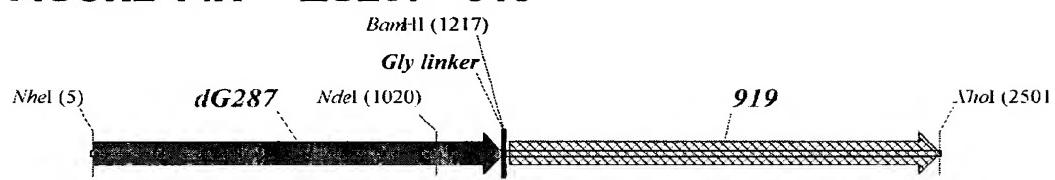
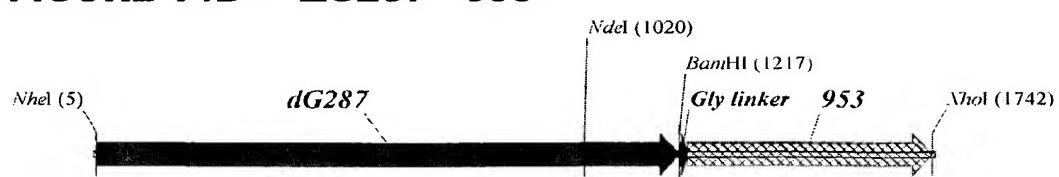


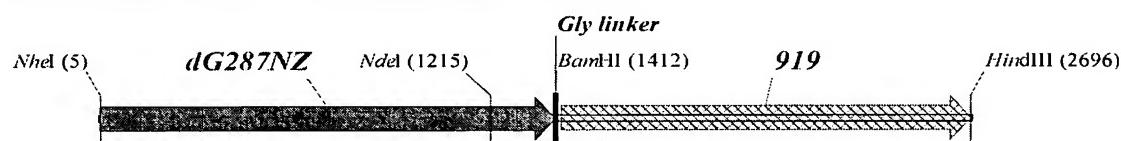
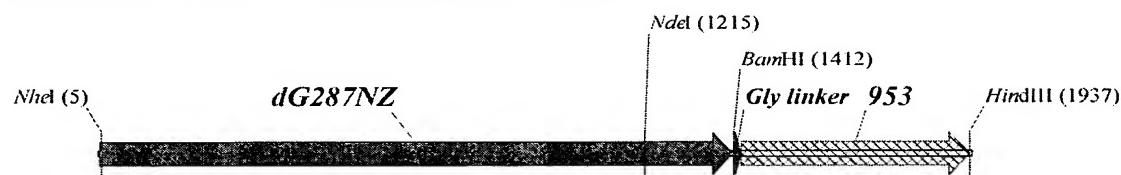
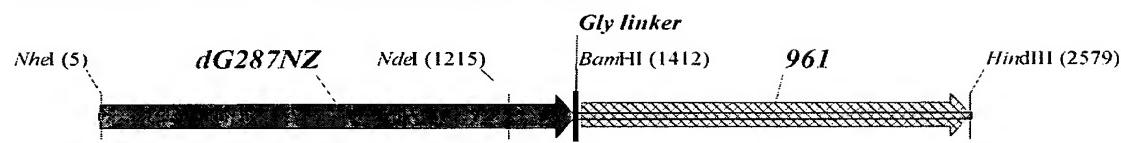
Leader Peptide

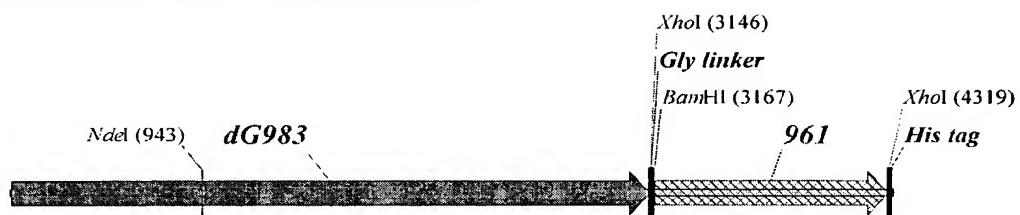
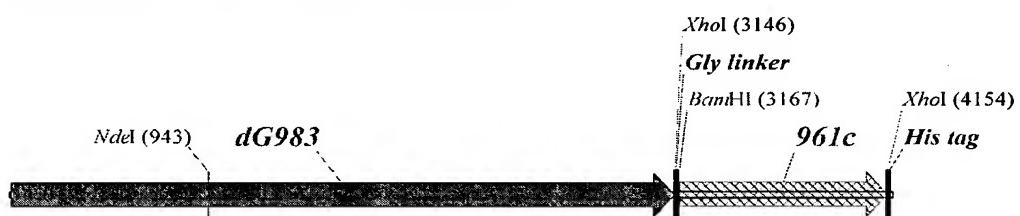
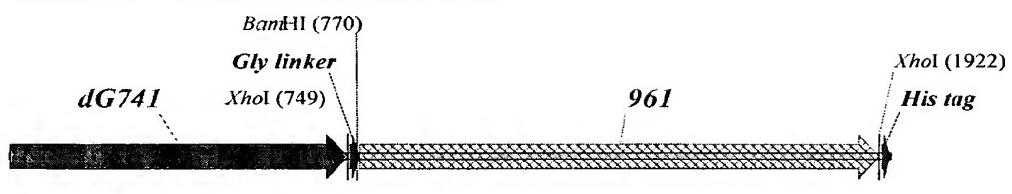
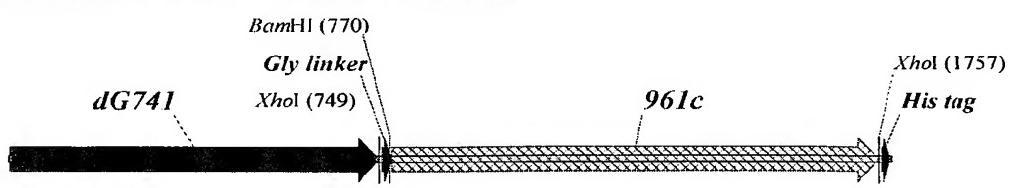
Region present in 2996,  
not in MC58

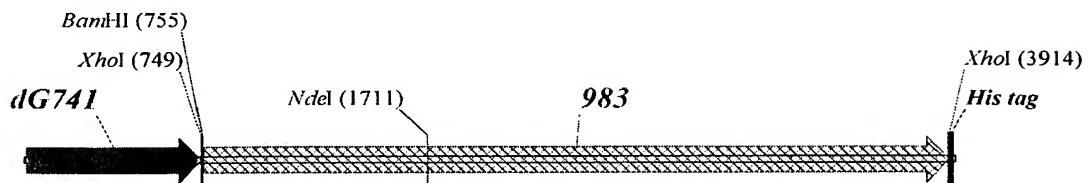
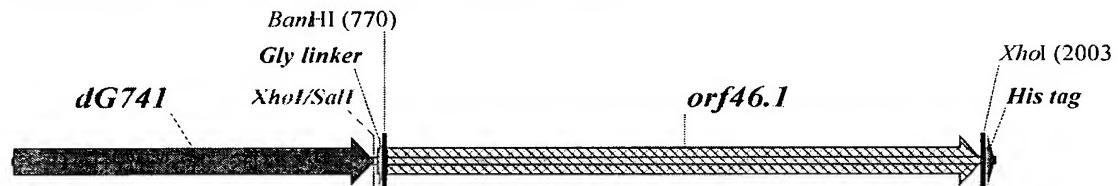
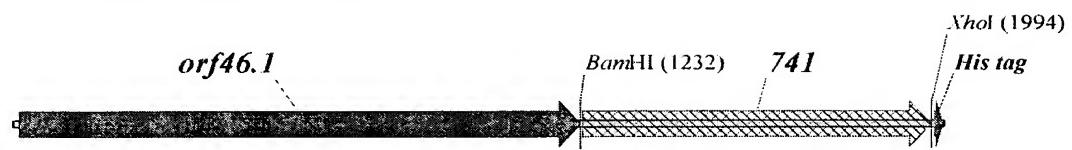
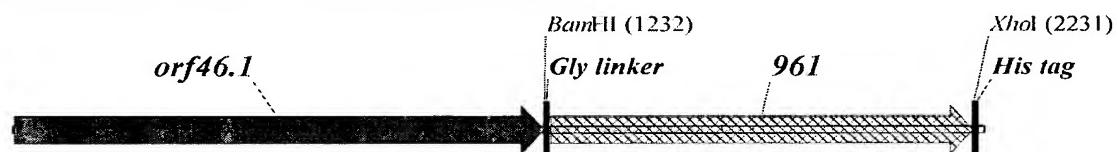
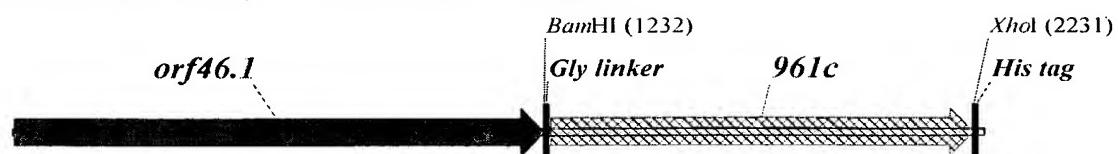
Coiled-coiled segment

Membrane anchor

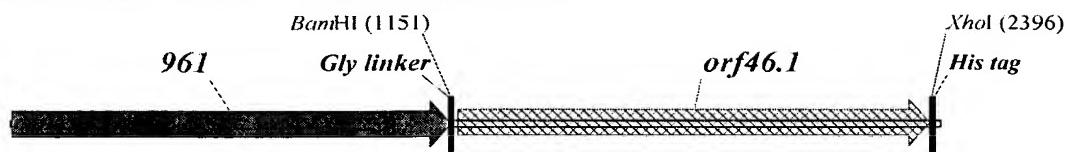
**FIGURE 13****FIGURE 14****FIGURE 14A — ΔG287—919****FIGURE 14B — ΔG287—953**

**FIGURE 14C —  $\Delta G287-961$** **FIGURE 14D —  $\Delta G287NZ-919$** **FIGURE 14E —  $\Delta G287NZ-953$** **FIGURE 14F —  $\Delta G287NZ-961$** **FIGURE 14G —  $\Delta G983-ORF46.1$** 

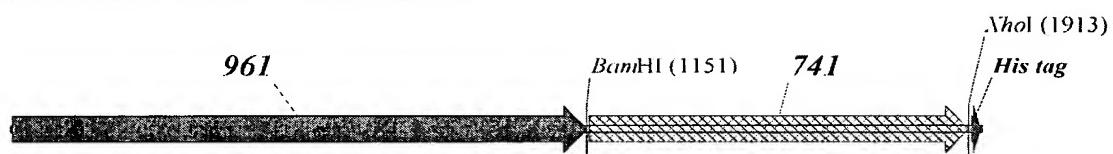
**FIGURE 14H —  $\Delta G983\text{-}741$** **FIGURE 14I —  $\Delta G983\text{-}961$** **FIGURE 14J —  $\Delta G983\text{-}961c$** **FIGURE 14K —  $\Delta G741\text{-}961$** **FIGURE 14L —  $\Delta G741\text{-}961c$** 

**FIGURE 14M — ΔG741-983****FIGURE 14N — ΔG741-ORF46.1****FIGURE 14O — ORF46.1-741****FIGURE 14P — ORF46.1-961****FIGURE 14Q — ORF46.1—961c**

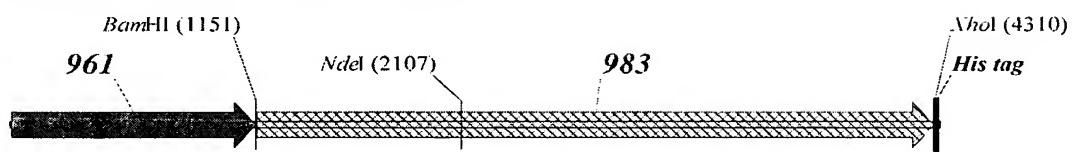
**FIGURE 14R — 961-ORF46.1**



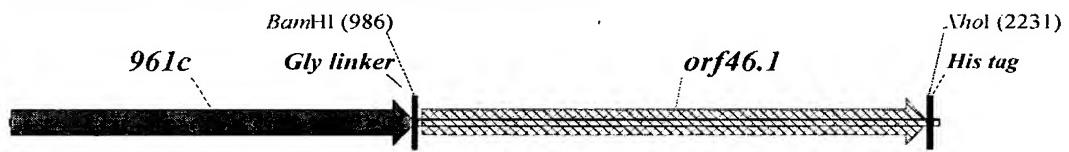
**FIGURE 14S — 961-741**



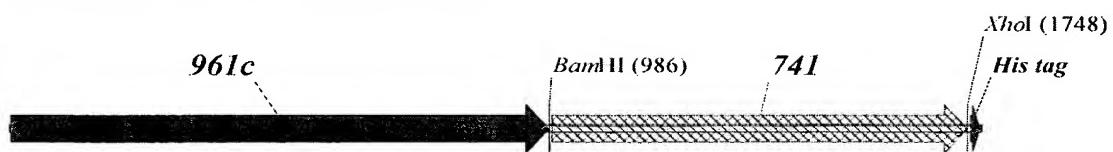
**FIGURE 14T — 961-983**

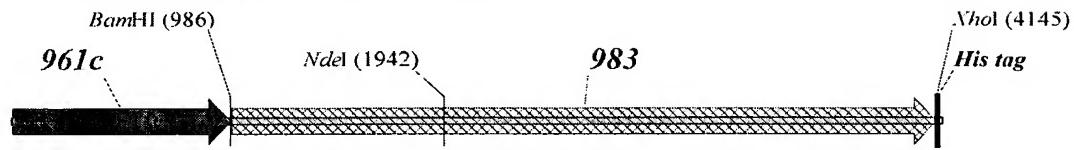
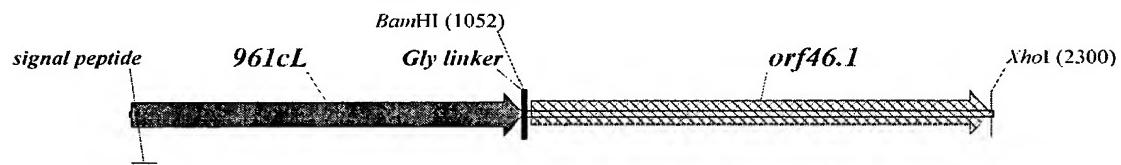
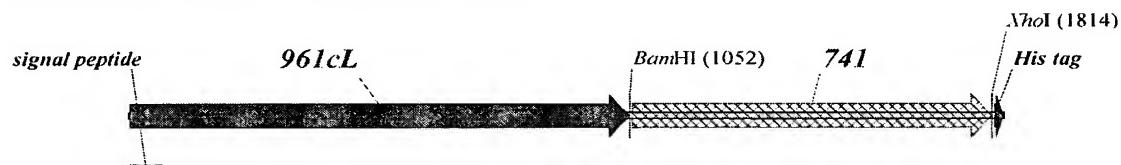
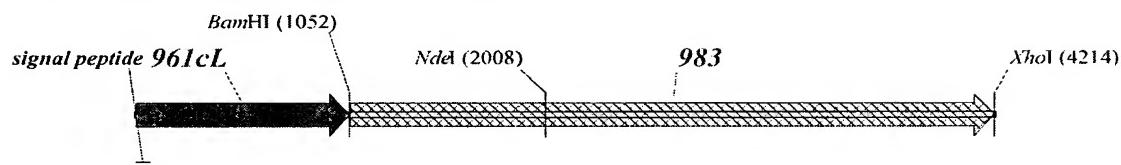


**FIGURE 14U — 961c-ORF46.1**



**FIGURE 14V — 961c-741**



**FIGURE 14W — 961c-983****FIGURE 14X — 961cL-ORF46.1****FIGURE 14Y — 961cL-741****FIGURE 14Z — 961cL-983**

## REFERENCES CITED IN THE DESCRIPTION

*This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.*

## Patent documents cited in the description

- WO 9924578 A [0002] [0004] [0004] [0005] [0005] [0149] [0178]
- WO 9936544 A [0002] [0004] [0004] [0005] [0005] [0178]
- WO 9957280 A [0002] [0004] [0004] [0005] [0005] [0102] [0167] [0173] [0178]
- WO 0022430 A [0002]
- WO 0066791 A [0006]
- WO 0066741 A [0008] [0035] [0037] [0053] [0074] [0086] [0101] [0130] [0158] [0172] [0194] [0279]
- WO 0071574 A [0009]
- WO 0104316 A [0009]
- WO 9955873 A [0148]

## Non-patent literature cited in the description

- TETTELIN et al. *Science*, 2000, vol. 287, 1809-1815 [0006]
- URSINUS ; HOLTJE. *J.Bact.*, 1994, vol. 176, 338-343 [0111]
- HARZ. *Anal. Biochem.*, 1990, vol. 190, 120-128 [0114]
- GLAUNER. *Anal. Biochem.*, 1988, vol. 172, 451-464 [0117]
- THUNNISSEN et al. *Biochemistry*, 1995, vol. 34, 12729-12737 [0121]
- VAN ASSELT et al. *Structure Fold Des*, 1999, vol. 7, 1167-80 [0121]
- TERRAK et al. *Mol.Microbiol.*, 1999, vol. 34, 350-64 [0124]
- VELATI BELLINI et al. *J. Biotechnol.*, 1991, vol. 18, 177-192 [0140]
- ROPP ; NICHOLAS. *J. Bact.*, 1997, vol. 179, 2783-2787 [0190]
- KRAFT. *J. Bact.*, 1998, vol. 180, 3441-3447 [0191]
- THALLER et al. *Microbiology*, 1994, vol. 140, 1341-1350 [0248]
- HOICZYK et al. *EMBO J*, 2000, vol. 19, 5989-99 [0272]
- EVANS. *Infect.Immun.*, 1974, vol. 10, 1010-1017 [0283]
- FILIP. *J.Bact.*, 1973, vol. 115, 717-722 [0307]
- DAVIES. *J.Immunol.Meth.*, 1990, vol. 143, 215-225 [0307]